

Open Data en Business Intelligence

Een exploratief onderzoek naar de bekendheid met, en de toepassing van, Open Data door Business Intelligence binnen private organisaties.

Auteur

J.W.A. van Poppel

Studentnummer

851226781

Datum

22-11-2015

Open Data en Business Intelligence – Een exploratief onderzoek naar de bekendheid met, en de toepassing van, Open Data door Business Intelligence binnen private organisaties.

Open Data and Business Intelligence – An explorative research on the awareness and use of Open Data by Business Intelligence within private organizations.

Open Universiteit, faculteit Management, Science & Technology
Masteropleiding Business Process Management & IT

Eerste begeleider	Dhr. prof. dr. ir. R.W. Helms
Tweede begeleider	Mevr. dr. A.D. Counotte-Potman
Examinator	Dhr. prof. dr. ir. R.W. Helms
Cursuscode	T9232B
Publicatiedatum en plaats	22 november 2015, Oisterwijk

Voorwoord

Na de succesvolle afronding van de premaster Managementwetenschappen en de inhoudelijke modules van de opleiding Business Process Management & IT ben ik in december 2014 gestart met mijn afstudeeronderzoek. Na een jaar hard werken als onderzoeker ben ik dan eindelijk klaar. Met trots presenteer ik dit eindproduct.

De resultaten van het uitgevoerde onderzoek zijn niet het werk van één persoon. Verschillende mensen hebben bijgedragen aan het resultaat. Zonder hun hulp was dit niet gelukt. Daarom wil ik via deze weg een aantal mensen bedanken. Allereerst mijn begeleider Remko Helms. Ik wil je bedanken voor de constructieve samenwerking. Bij twijfel of onduidelijkheid wist je me altijd de juiste energie te geven om verder te gaan. Je voorzag me van nieuwe inzichten in de vorm van artikelen of pushte me om zelf die inzichten te vormen. Daarbij promootte je verschillende tools als Mendeley en NVivo om het afstudeerproces aangenamer en kwalitatief beter te maken. Dank gaat ook uit naar mijn tweede begeleider, Anda Counotte-Potman. Daar waar het afstudeerproces onduidelijk was, zorgde je in jouw rol van afstudeercoördinator Informatica voor helder inzicht. Als meelezer heb je me zeer geholpen met het reviewen van mijn mijlpaaldocumenten en thesis. De respondenten, zowel vanuit de survey als uit de casestudy, zijn van onnoemelijk belang geweest. Ik wil jullie bedanken voor jullie bijdrage.

Een speciale dank aan mijn vrouw Noortje en onze kinderen Guus en Bram. Jullie steun en begrip is de afgelopen vier jaar ontzettend belangrijk geweest. Zonder jullie was ik niet zo ver gekomen.

Joost van Poppel

Oisterwijk, 22-11-2015

Inhoudsopgave

Samenvatting	7
1. Inleiding.....	9
1.1. Aanleiding	9
1.2. Doelstelling	9
1.3. Aspecten van Business Intelligence	9
1.4. Hoofdvraag en deelvragen	10
1.5. Relevantie	10
1.6. Aanpak onderzoek	11
1.7. Leeswijzer	12
2. Onderzoeksstrategie	13
2.1. Literatuurstudie	13
2.1.1. Kaders.....	13
2.1.2. Proces van uitvoering.....	13
2.1.3. Trefwoorden definiëren	14
2.1.4. Zoekopdrachten uitvoeren.....	14
2.1.5. Literatuur opslaan	15
2.1.6. Literatuur beoordelen	15
2.1.7. Concept overzicht.....	15
2.2. Empirisch onderzoek	16
2.2.1. Onderzoeksstrategie	16
2.2.2. Bronnen en middelen.....	17
2.2.3. Betrouwbaarheid, validiteit en ethiek	19
2.2.4. Wijze van analyseren.....	21
2.2.5. Verantwoording survey en casestudy	22
3. Literatuurstudie.....	23
3.1. Literatuur Open Data en Business Intelligence	23
3.2. L1 – Wat is Open Data?	23
3.2.1. Open Movements.....	23
3.2.2. Data, gegevens en informatie	24
3.2.3. Definitie van Open Data	25
3.2.4. Gerelateerde concepten	26

3.2.5.	Categorieën Open Data	27
3.3.	L2 – Wie stelt Open Data beschikbaar?	28
3.3.1.	Bronnen	28
3.3.2.	Redenen van openstelling	28
3.3.3.	Promotie gebruik Open Data	29
3.4.	L3 – Wat zijn de kenmerken van Open Data?	30
3.4.1.	Kenmerken van Open Data	30
3.4.2.	Mate van hergebruik	31
3.4.3.	Linked Open Data	31
3.4.4.	Licentievrij	32
3.5.	L4 – Hoe kan Open Data worden toegepast?	33
3.5.1.	Websites & Apps	33
3.5.2.	Bedrijfsmodel	33
3.6.	L5 – Wat is Business Intelligence?	34
3.7.	L6 – Hoe ziet een Business Intelligence architectuur er uit?	35
3.7.1.	Databronnen en verplaatsing van data	35
3.7.2.	Data warehousing	35
3.7.3.	De middenlaag	36
3.7.4.	Front-end toepassingen	36
3.8.	L7 – Welke functies kent Business Intelligence?	37
3.9.	Eindresultaat literatuurstudie	38
4.	Onderzoeksresultaten	39
4.1.	Survey – Beantwoording E1 en E2	39
4.1.1.	Respons	39
4.1.2.	Achtergrond respondenten	40
4.1.3.	Achtergrond Business Intelligence	41
4.1.4.	Is Business Intelligence op de hoogte van het fenomeen Open Data?	42
4.1.5.	Maakt Business Intelligence gebruik van Open Data?	42
4.2.	Meervoudige casestudy – Beantwoording E3 en E4	44
4.2.1.	Cases	44
4.2.2.	Algemene kenmerken van de cases	44
4.2.3.	Case: Achmea	44
4.2.4.	Case: Enxsis	51

4.2.5.	Case: Anoniem.....	57
4.2.6.	Overeenkomsten en verschillen bij de cases	59
4.2.7.	Overeenkomsten en verschillen per case geplot naar het referentiemodel	60
5.	Conclusies en aanbevelingen	61
5.1.	Conclusies literatuurstudie	61
5.2.	Conclusies empirisch onderzoek	63
5.3.	Eindconclusie	66
5.4.	Aanbevelingen	67
5.5.	Limitaties van het onderzoek	68
6.	Procesreflectie.....	69
	Referenties	70
	Bijlagen.....	73
	Bijlage A – Contacttemplates	74
	Bijlage B – Survey navigatiepad.....	76
	Bijlage C – Coderingstabellen survey	77
	Bijlage D – Protocol casestudy	83
	Bijlage E – Correspondentie auteurs Open Data artikelen.....	85
	Bijlage F – Referentiemodel.....	89
	Bijlage G – Operationalisering.....	91
	Bijlage H – Survey resultaten.....	92
	Bijlage I – Overzicht documenten documentatiestudie	107
	Bijlage J – Boomstructuur NVivo	108

Samenvatting

Ter ondersteuning van operationele, tactische en strategische besluitvorming wordt binnen organisaties vaak gebruik gemaakt van Business Intelligence. Business Intelligence heeft als doel om kenniswerkers te voorzien van besluitvorming ondersteunende technologieën. Daarbij wordt gebruik gemaakt van interne data, de data die de organisatie zelf maakt, en externe data, data die buiten de organisatie wordt gecreëerd. Een nieuw te ontsluiten externe bron van data is Open Data. Open Data is data die vrij gebruikt kan worden, hergebruikt kan worden en opnieuw verspreid kan worden door iedereen – onderworpen enkel, in het uiterste geval, aan de voorwaarde dat men de bron vermeldt en de data ook weer deelt. Open Data kan bestaande diensten overbodig maken, nieuwe inzichten bieden in de bedrijfsvoering, kernactiviteiten van organisaties verbeteren, kosten verlagen en de eigenschappen van de externe omgeving inzichtelijk maken waardoor een grotere economische waarde behaald kan worden. Business Intelligence, dat zich sterk leent voor de ontsluiting van interne en externe bronnen, kan een belangrijke rol spelen in het gebruik van Open Data. Met overheden als hoofdproducent van Open Data is er geen inzicht in de huidige toepassing van Open Data door Business Intelligence bij private organisaties. De doelstelling van dit exploratieve onderzoek is om hierin inzicht te creëren.

Daartoe is de volgende hoofdvraag geformuleerd: ***In hoeverre is Business Intelligence binnen private organisaties zich bewust van het bestaan van Open Data, en op welke manier passen ze Open Data toe?***

Om een antwoord te formuleren op de hoofdvraag is een literatuurstudie en een empirisch onderzoek uitgevoerd.

Omdat er geen documentatie beschikbaar is over de toepassing van Open Data door Business Intelligence is tijdens de literatuurstudie een splitsing gemaakt. Enerzijds is onderzocht wat Open Data is, wie Open Data beschikbaar stelt, welke kenmerken Open Data hebben en hoe Open Data kan worden toegepast. Anderzijds is onderzocht wat Business Intelligence is, hoe een typische Business Intelligence architectuur er uit ziet en welke functies de verschillende Business Intelligence technologieën kennen. Daaruit is een samenvattend referentiemodel ontstaan van waaruit de operationalisering heeft plaatsgevonden voor de empirische studie.

Het empirisch onderzoek is tweeledig uitgevoerd. Om antwoord te geven op de vraag of Business Intelligence zich bewust is van het bestaan van Open Data en of Business Intelligence daadwerkelijk gebruik maakt van Open Data is een survey uitgezet aan de top 500 organisaties van Nederland. Op basis van de reactie van 67 respondenten zijn conclusies getrokken. Om antwoord te geven op de vraag op welke manier Business Intelligence Open Data toepast, en welke problemen zij daarbij ondervinden, is een meervoudige casestudy uitgevoerd bij Achmea, Enexis en een anonieme organisatie. Op basis van 14 interviews, verdeeld over de 3 casussen, en een documentatiestudie, zijn conclusies getrokken.

Uit de resultaten van de survey blijkt dat 78%, 52, van de respondenten op de hoogte is van het bestaan van het fenomeen Open Data. Daarvan wordt door 33%, 17, van de organisaties daadwerkelijk gebruik gemaakt van Open Data. Organisaties die geen gebruik maken van Open Data geven aan dat ze geen behoefte hebben aan Open Data, dat ze te intern gefocust zijn en dat er geen revelante data beschikbaar is. Uit de resultaten van de meervoudige casestudy blijkt dat de Business Intelligence organisaties veel van elkaar weghebben als het gaat om e aandachtsgebieden, architectuur en functies. Er zijn duidelijke overeenkomsten en verschillen waarneembaar in de toepassing van Open Data.

De beantwoording van de hoofvraag luidt: Het aantal Business Intelligence organisaties dat zich bewust is van het bestaan van Open Data is hoog terwijl een derde van de organisaties er daadwerkelijk gebruik van maakt. De toepassing van Open Data, het initiatief komt vaak niet van Business Intelligence zelf maar vanuit de business, vindt ad hoc of structureel plaats. Bij ad hoc toepassing worden door Business Intelligence stand alone applicaties en Self Service omgevingen geboden terwijl bij structurele toepassing de Open Data als externe bron wordt ontsloten en wordt geladen in het datawarehouse waarna het verder beschikbaar wordt gesteld aan de middenlaag en de front-ends. De toepassing van Open Data heeft op verschillende aandachtgebieden, architectuuronderdelen en functies van Business Intelligence invloed. Kenmerkend bij de toepassing van Open Data is dat vooral gebruik wordt gemaakt van datasets afkomstig van de overheid. Bij de toepassing worden verschillende categorieën Open Data aangeboord en houdt men rekening met het belang van de verschillende kenmerken. Bij de toepassing van Open Data wordt met een aantal problemen rekening gehouden: De kwaliteit is vaak onder de maat, er is te veel Open Data beschikbaar, de gewenste Open Data is niet beschikbaar, Open Data is niet tijdig beschikbaar en Open Data is niet gemakkelijk te koppelen met eigen data. Bij de toepassing van Open Data is een mate van volwassenheid zichtbaar, variërend van het niet gebruiken van Open Data tot het gebruikmaken én publiceren van Open Data.

1. Inleiding

1.1. Aanleiding

Ter ondersteuning van operationele, tactische en strategische besluitvorming wordt binnen organisaties vaak gebruik gemaakt van Business Intelligence. Het heeft als doel om kenniswerkers als managers, analisten en planners te voorzien van besluitvorming ondersteunende technologieën (Chaudhuri, Dayal, & Narasayya, 2011). Business Intelligence wordt gedefinieerd als “een verzameling technieken en hulpmiddelen voor de omzetting van ruwe data naar betekenisvolle en bruikbare informatie voor business analyse mogelijkheden. Business Intelligence technologieën zijn in staat om grote hoeveelheden data te verwerken ter identificatie, ter ontwikkeling en ter ontwerp van nieuwe strategische kansen.” (Rud, 2009). Business Intelligence is het meest effectief wanneer het data uit de omgeving van de organisatie combineert met data die door de organisatie zelf wordt gecreëerd. Deze combinatie van verschillende interne en externe data zorgt voor een mate van intelligentie die niet bereikt kan worden door het gebruiken van een enkele set data (Bures, Otcenaskova & Jasikova, 2012).

Een nieuw te ontsluiten externe databron bij Business Intelligence is Open Data. “Open Data is data die vrij gebruikt kan worden, hergebruikt kan worden en opnieuw verspreid kan worden door iedereen – onderworpen enkel, in het uiterste geval, aan de voorwaarde dat men de bron vermeldt en de data ook weer deelt.” (Borglund & Engvall, 2014; Lindman, 2014; Ubaldi, 2013).

Open Data vormt een nieuwe externe databron voor Business Intelligence. Het kan bestaande betaalde diensten overbodig maken, het kan nieuwe inzichten bieden in de bedrijfsvoering, kernactiviteiten van organisaties verbeteren (Kaasenbrood & Zuiderwijk, 2015), kosten verlagen en de eigenschappen van de externe omgeving inzichtelijk maken waardoor een grotere economische waarde behaald kan worden.

Om een antwoord te kunnen geven op de vraag in hoeverre de Business Intelligence afdeling/discipline binnen private organisaties zich bewust zijn van het bestaan van Open Data, en op welke manier ze deze toepassen, is dit onderzoek uitgevoerd.

1.2. Doelstelling

In de literatuur blijkt het verband tussen Open Data met Business Intelligence onderbelicht. De doelstelling van het onderzoek is om enerzijds inzichtelijk te maken of Business Intelligence van het bestaan van Open Data afweet en anderzijds inzichtelijk te maken hoe Business Intelligence gebruik maakt van Open Data, en welke problemen daarbij worden ervaren.

1.3. Aspecten van Business Intelligence

In de definitie van Business Intelligence worden drie aspecten onderscheiden (BI Award, 2015). Er is een onderscheid te maken naar Business Intelligence als proces, technologie en fenomeen. Het onderzoek richt zich op de Business Intelligence *discipline* binnen organisaties. Vaak ondergebracht in een team, afdeling of organisatie.

1.4. Hoofdvraag en deelvragen

Voor het bereiken van het doel van het onderzoek is de volgende hoofdvraag geformuleerd:

In hoeverre is Business Intelligence binnen private organisaties zich bewust van het bestaan van Open Data, en op welke manier passen ze Open Data toe?

Ter beantwoording van de hoofdvraag worden een aantal deelvragen beantwoord. In de volgende twee tabellen (1 en 2) worden respectievelijk de onderzoeksvragen getoond die vanuit de literatuur en de empirie worden beantwoord. Ter identificatie is een L en een E gebruikt.

ID	Deelvraag
L1	Wat is Open Data?
L2	Wie stelt Open Data beschikbaar?
L3	Welke kenmerken hebben Open Data?
L4	Hoe kan Open Data worden toegepast?
L5	Wat is Business Intelligence?
L6	Hoe ziet een typische Business Intelligence architectuur er uit?
L7	Welke functies hebben de verschillende Business Intelligence technologieën?

Tabel 1 - Deelvragen te beantwoorden uit literatuurstudie

ID	Deelvraag
E1	Is Business Intelligence op de hoogte van het bestaan van Open Data?
E2	Maakt Business Intelligence gebruik van Open Data?
E3	Op welke manier past Business Intelligence Open Data toe?
E4	Welke problemen worden ondervonden bij het gebruik van Open Data door Business Intelligence?

Tabel 2 - Deelvragen te beantwoorden uit empirisch onderzoek

1.5. Relevantie

In de wetenschappelijke literatuur is veel geschreven over Open Data. De nadruk ligt voornamelijk op het beschikbaar stellen van Open Data. In beperkte mate wordt het gebruik of de toepassing beschreven. Aan een verdere verbijzondering van het gebruik van Open Data, bij Business Intelligence, is zelfs geen enkel artikel gewijd. Door uitvoering van het onderzoek wordt een theoretische leemte gevuld over de toepassing van Open Data door Business Intelligence.

De uitkomsten van het onderzoek bieden ook praktische toegevoegde waarde voor Business Intelligence. Naast inzichten in soorten, kenmerken, vormen, bronnen en toepassingsmogelijkheden van Open Data, biedt het onderzoek ook inzicht in hoe Business Intelligence Open Data toepast en welke problemen daarbij ondervonden worden.

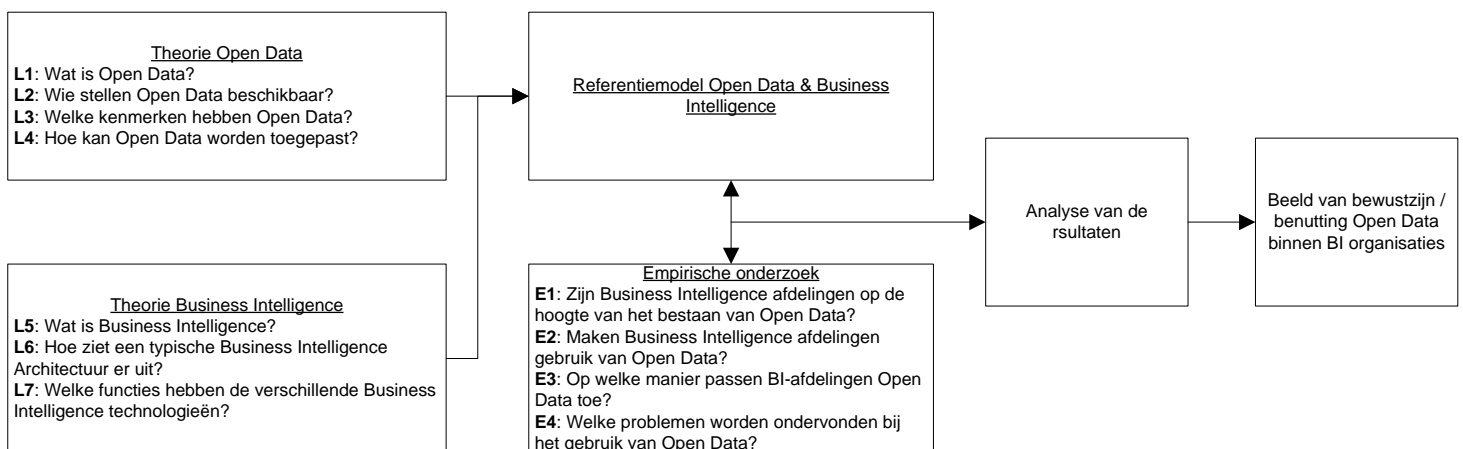
1.6. Aanpak onderzoek

Het onderzoek is benaderd als een onderzoeksproject met daarin verschillende fasen. Na het opstellen van een voorlopige opdrachtformulering is de literatuurstudie uitgevoerd. Deze literatuurstudie was er op gericht om de deelvragen L1 t/m L7 te beantwoorden. Deelvragen L1 t/m L4 gaan over Open Data. Deelvragen L5 t/m L7 gaan over Business Intelligence. Deze specifieke splitsing naar Open Data en Business Intelligence was noodzakelijk omdat er tijdens het opstellen van de voorlopige opdrachtformulering geen literatuur werd gevonden over de combinatie. Het product van de literatuurstudie betreft een referentiemodel over Open Data en een referentiemodel over Business Intelligence. Na afronding van de literatuurstudie is de opdrachtformulering definitief gemaakt.

Om antwoord te kunnen geven op de empirische vragen E1 t/m E4 is een exploratief onderzoek uitgevoerd. Allereerst is een survey uitgezet aan de Business Intelligence discipline/afdeling van de top 500 private organisaties van Nederland. Gevolgd door een meervoudige casestudy bij een drietal private organisaties om de empirische vragen E3 en E4 te beantwoorden.

Na de dataverzameling zijn de resultaten geanalyseerd, conclusies getrokken en aanbevelingen geformuleerd. Dit document verschaft daarmee een beeld van het bewustzijn en gebruik van Open Data door Business Intelligence disciplines/afdelingen.

Figuur 1 toont een visueel overzicht van het onderzoek in de vorm van een onderzoeksmodel.



Figuur 1 - Onderzoeksmodel

1.7. Leeswijzer

- Hoofdstuk 1: Betreft de inleiding met daarin beschreven het centrale onderzoeksthema dat onderzocht is. Onderdelen betreffen de aanleiding en doelstelling van het onderzoek, de hoofd- en deelvragen, een beschrijving van de relevantie en een overzicht van de onderzoeksaanpak.
- Hoofdstuk 2: Hier wordt de onderzoeksaanpak in detail uiteengezet. In de eerste plaats zal de aanpak van de literatuurstudie worden weergegeven gevolgd door een beschrijving van de aanpak van het empirische onderzoek.
- Hoofdstuk 3: Bevat de resultaten van de literatuurstudie, de beantwoording van de deelvragen L1 t/m L7.
- Hoofdstuk 4: Bevat de resultaten van het empirische onderzoek. De deelvragen E1 en E2 worden beantwoord op basis van de survey. De deelvragen E3 en E4 worden beantwoord vanuit de meervoudige casestudy.
- Hoofdstuk 5: De conclusies en aanbevelingen. Hier wordt een algemene conclusie gegeven op de deelvragen uit de literatuurstudie en de empirische studie. Daarbij zijn aanbevelingen gedaan voor vervolgonderzoek.
- Hoofdstuk 6: Bevat de procesreflectie. Er wordt hier gereflecteerd op het uitgevoerde onderzoek. Wat waren de tips en wat waren de tops.

2. Onderzoeksstrategie

Om de centrale vraag te kunnen beantwoorden, moeten de deelvragen beantwoord worden. De beantwoording van de deelvragen heeft plaatsgevonden door de uitvoering van een literatuurstudie, gevolgd door een empirisch onderzoek. Vanuit de literatuurstudie is een samenvattend model over Open Data en Business Intelligence ontwikkeld dat ter referentie heeft gediend bij het empirisch onderzoek. In dit hoofdstuk vindt de methodische verantwoording plaats.

2.1. Literatuurstudie

“Het hoofddoel van de literatuurstudie is het presenteren van een kritisch theoretisch kader om het onderzoek in een bredere context te plaatsen en om te laten zien op welke wijze het onderzoek bijdraagt aan het werk dat al eerder is verricht over dit onderzoeksthema en topic.” (Hofstee & Kusters, 2013). Om richting te geven aan de literatuurstudie worden de zeven deelvragen beantwoord (L1 t/m L7).

2.1.1. Kaders

2.1.1.1. Literatuurbasis

Naast wetenschappelijke literatuur zijn op het Internet veel andere bronnen beschikbaar. Het is niet gemakkelijk vast te stellen wat de kwaliteit van een bepaalde bron is. In dit onderzoek wordt de nadruk gelegd op wetenschappelijke literatuur. Bij de zoektocht naar literatuur ligt de primaire focus op peer-reviewed artikelen uit erkende internationale journals en conference proceedings. De literatuur is op het Internet gepubliceerd en naspeurbaar. Er zijn tijdens de literatuurstudie Engelse, Nederlandse en Duitse literatuurstukken geraadpleegd. De onderzoeker is met deze talen bekend.

2.1.1.2. Publicatietijdstip

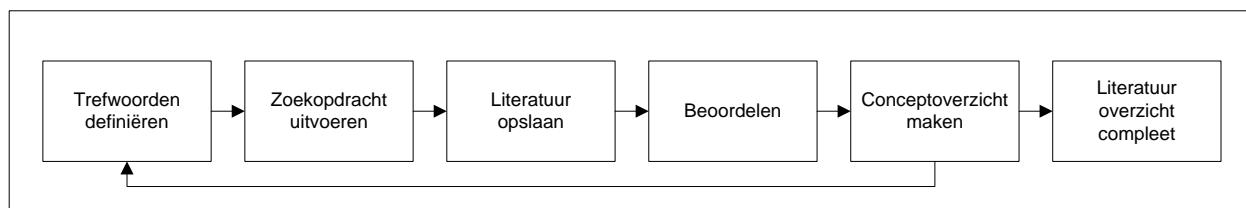
Bij de zoektocht naar literatuur is het publicatietijdstip buiten beschouwing gebleven. Het belang van het literatuurstuk staat centraal.

2.1.1.3. APA-stijl

Bij het verwijzen naar literatuur in dit document wordt door gebruikmaking van de APA-6 stijl verwezen naar de bron. Deze stijl is opgesteld door de American Psychological Association (APA), een organisatie voor academici in de Verenigde Staten.

2.1.2. Proces van uitvoering

Het proces van de uitvoering van de literatuurstudie is iteratief verlopen zoals weergegeven in figuur 2. Dit proces (Saunders, Lewis, Thornhill, Booi, & Verckens, 2011) zorgt niet alleen voor een filtering op de bruikbare literatuur, maar zorgt er ook voor dat in opvolgende iteraties naar meer relevante literatuur gezocht kan worden. Gedurende de literatuurstudie zijn een viertal iteraties doorlopen.



Figuur 2 - Proces van uitvoering van literatuurstudie

2.1.3. Trefwoorden definiëren

Per iteratie zijn een aantal trefwoorden gedefinieerd in de zoektocht naar relevante artikelen. In tabel 3 wordt een weergave gegeven van de gebruikte trefwoorden per iteratie. Met de trefwoorden tussen aanhalingstekens is gezocht op basis van AND-operatoren.

Iteratie	Gebruikte trefwoorden
1	"Open Data Business Intelligence" "Open Data Utilization in Business Intelligence" "Open Data" "Open Data Usage" "Open Data Open Source"
2	"Open Government Data Business Intelligence" "Public Sector Information Business Intelligence" "Open Government Data" "Public Sector Information"
3	"Business Intelligence" "Business Intelligence Architecture"
4	"Business Intelligence Systems" "Business Intelligence Functions"

Tabel 3 - Gebruikte trefwoorden literatuurstudie

2.1.4. Zoekopdrachten uitvoeren

De zoekopdrachten zijn uitgevoerd door gebruik te maken van de Digitale Bibliotheek van de Open Universiteit. De Digitale Bibliotheek is exclusief toegankelijk voor studenten en door het gebruikmaken hiervan hoeven geen artikelkosten te worden betaald. Dit medium bevat functionaliteit waarmee men direct in verschillende informatiebestanden en zoekmachines kan zoeken. Hierbij kan als filter-optie worden meegegeven dat men alleen peer-reviewed artikelen wil zien als resultaat. De bronnen die gekoppeld zijn aan de Digitale Bibliotheek worden in tabel 4 opgesomd.

Academic Search Elite (EBSCO)	Google Scholar	ScienceDirect (Elsevier)
ACM Digital Library	IEEE Digital Library	SpringerLink
Cambridge University Press	Lecture Notes in Computer Science	Taylor & Francis Group
EBSCO Host	Oxford Journals	Web of Science
Emerald – Management Plus	SAG Journals Online	Wiley Online Library
ERIC (EBSCO)	Science	WorldCat

Tabel 4 - Bronnen te benaderen via Digitale Bibliotheek

2.1.5. Literatuur opslaan

Ter documentopslag en verdere ondersteuning van het literatuurproces is gebruik gemaakt van een 'Referencemanager'. Hiervoor is gebruik gemaakt van de applicatie Mendeley. Mendeley maakt het eenvoudig om wetenschappelijke artikelen te beheren en te bestuderen. Daarbij beschikt de applicatie over een Microsoft Word plug-in waarmee eenvoudig de referenties toegevoegd kunnen worden in een bibliografie. Alle gevonden artikelen zijn in Mendeley opgeslagen.

2.1.6. Literatuur beoordelen

Op basis van de samenvatting, de conclusie en de aanbevelingen is beoordeeld of artikelen wel al dan niet geschikt waren voor verdere bestuderingen. Waar artikelen niet relevant bleken, zijn deze verwijderd uit Mendeley.

2.1.7. Concept overzicht

Waar artikelen wel relevant bleken, zijn deze bestudeerd, beoordeeld, en is er een samenvatting van gemaakt.

2.2. Empirisch onderzoek

Op basis van het gegenereerde theoretisch kader uit de literatuurstudie wordt in het empirisch onderzoek gezocht naar het antwoord op de vier empirische vragen (E1 t/m E4). De empirische vragen E1 en E2 zijn beantwoord door gebruikmaking van een survey, de empirische vragen E3 en E4 zijn beantwoord door uitvoering van een casestudy.

2.2.1. Onderzoeksstrategie

Omdat de toepassing van Open Data door Business Intelligence onderbelicht is in de literatuur, is het onderzoek dat uitgevoerd is te bestempelen als verkennend of exploratief. De deelvragen zijn onderzocht door gebruik te maken van twee kwalitatieve onderzoeksstrategieën. Dit wordt in de literatuur een multimethodisch onderzoek genoemd (Saunders et al., 2011). Bij multimethodisch onderzoek wordt dus óf alleen kwalitatief óf alleen kwantitatief onderzoek verricht (Tashakkori & Teddlie, 2003).

Bewustzijn van het bestaan van Open Data

De eerste twee empirische vragen luiden:

- E1: Is Business Intelligence op de hoogte van het bestaan van Open Data?
- E2: Maakt Business Intelligence gebruik van Open Data?

Om antwoord te kunnen geven op deze vragen is een survey ingezet. Voordeel van de inzet van de survey is dat deze ook gebruikt kan worden ter werving van respondenten voor de casestudy ter beantwoording van de vragen E3 en E4. De uitgezette survey is (semi-)kwalitatief van aard. "De kwalitatieve survey is een onderzoek naar de verscheidenheid in aard of betekenis van een of meer gedragingen, cognities of attributen binnen een populatie." (Jansen, 2005). Jansen (2005) geeft ook kort het verschil aan tussen de kwalitatieve en de kwantitatieve survey. De kwalitatieve survey is een onderzoek naar verscheidenheid en de kwantitatieve of statistische survey is een onderzoek naar frequentie. De survey kan een generaliserend beeld scheppen over de stand van zaken bij Business Intelligence in de private markt in Nederland.

Toepassing van Open Data

De laatste twee empirische vragen luiden:

- E3: Op welke manier past Business Intelligence Open Data toe?
- E4: Welke problemen worden ondervonden bij het gebruik van Open Data door Business Intelligence?

Deze vragen zijn beantwoord door inzet van een meervoudige casestudy. De casestudy is bij uitstek de onderzoeksstrategie voor een verkennend onderzoek. De casestudy is meervoudig uit noodzaak om te bepalen of de resultaten van de eerdere cases ook voorkomen in andere cases, en daardoor kunnen worden gegeneraliseerd (Jochems & Joosten, 2005). De casestudy is ingebed, en niet holistisch, omdat het onderzoek plaatsvindt bij Business Intelligence. Business Intelligence, vaak in de vorm van afdelingen, zijn subonderdelen van een grotere organisatie.

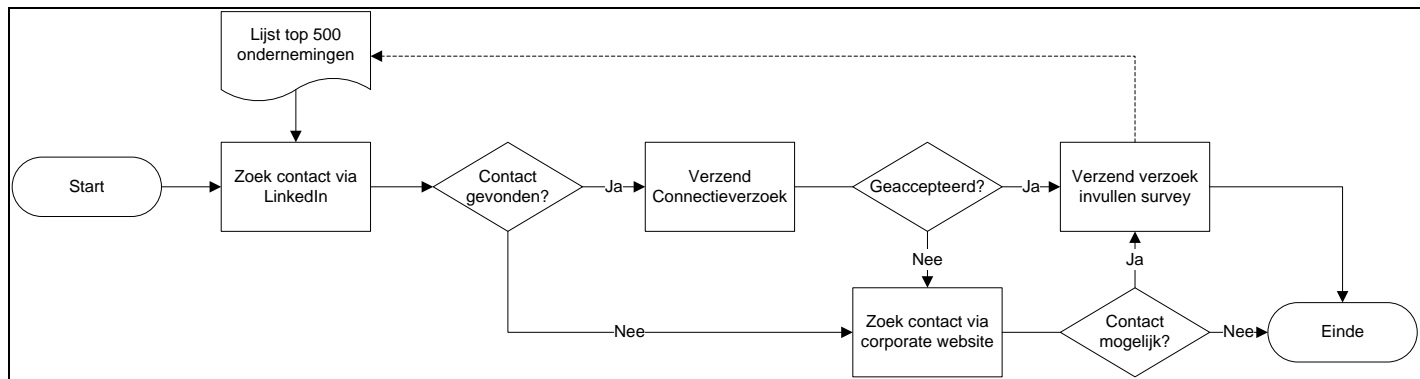
2.2.2. Bronnen en middelen

De empirische vragen E1 en E2 zijn beantwoord door gebruik te maken van een survey. De vragen E3 en E4 zijn beantwoord door middel van een casestudy. De gegevens die verzameld zijn, betreffen kwalitatieve primaire gegevens. De toegang tot gegevens is van cruciaal belang voor het empirisch onderzoek. De onderzoeker moet aan zijn of haar benodigde gegevens komen. Er kunnen verschillende problemen voordoen bij de werving van toegang tot de gegevens. Hieronder wordt verwoord hoe de survey en casestudy heeft plaatsgevonden.

Survey

Uit onderzoek blijkt dat organisaties die Business Intelligence inzetten vaak grote multinationals zijn met verschillende dochterondernemingen (Computerworld, 2006). De scope van het afstudeeronderzoek betreft de top 500 private organisaties in Nederland. Elsevier heeft in 2014 de namen van de 500 grootste organisaties gepubliceerd. Daarbij zijn verschillende gegevens gepubliceerd als de activiteit van de organisatie, de omzet/winst/verlies cijfers, het aantal medewerkers, etc. De lijst bestaat uit organisaties met een jaaromzet variërend tussen 231 miljoen en 353 miljard en het aantal medewerkers variërend tussen de 110 en 568.000. Voor de survey zijn de organisaties uit deze top 500 lijst benaderd.

De respondenten zijn benaderd volgens het processchema in figuur 3. Daarbij is van belang dat de verantwoordelijke vanuit Business Intelligence de survey invult omdat dat de persoon is die het beste overzicht heeft over Business Intelligence binnen de onderneming. Het processchema laat zien dat op basis van LinkedIn contact wordt gezocht met medewerkers Business Intelligence. Is een contact gevonden, dan wordt een connectieverzoek gestuurd naar het betreffende contact. Accepteert het contact het connectieverzoek, volgt een link voor de deelname aan de survey. Wordt er geen contact gevonden via LinkedIn, of wordt het connectieverzoek niet beantwoord, probeert de onderzoeker contact op te nemen via de corporate website van de onderneming. In bijlage A staan achtereenvolgens de templates voor het connectieverzoek, het verzoek tot invullen van de survey aan het LinkedIn contact en de mail voor benadering via de corporate website.



Figuur 3 - Procesflow werving respondenten voor survey

Voor de verzameling van data in de survey is gebruik gemaakt van het webscript 'LimeSurvey'. LimeSurvey is een open source script dat het mogelijk maakt om een online survey aan te maken. Dit script is geïnstalleerd op de website van de onderzoeker.

Op de vragenlijst van de survey staan 15 vragen, met een totale invulduur van ongeveer 5 minuten. Over het algemeen zal een korte duur van de survey bijdragen aan een hogere respons. De vragenlijst van de survey is opgenomen in bijlage B. De gebruikte coderingstabellen is opgenomen in bijlage C.

Casestudy

In het algemeen bestaan bij de uitvoering van casestudies drie waarnemingstechnieken. Dat betreft documentatiestudie, interviews en observatie. Om antwoord te kunnen geven op de empirische vragen, is gekozen om gebruik te maken van documentatiestudie, ongestructureerde en semigestructureerde interviews. Observatie is niet toegepast.

In de documentatiestudie zijn digitale documenten bestudeerd als notities, ontwerpen, gebruiksbeschrijvingen, etc. Daarbij zijn zowel primaire als secundaire documenten bekeken. Het voordeel van de documentatiestudie is dat de studie geen invloed heeft op personen, hij is non-reactief. Daarnaast is documentatiemateriaal veelal beschikbaar. Waar rekening mee gehouden is tijdens de uitvoering van de meervoudige casestudy is dat de documentatie niet zaligmakend is en dat er informatie in de documentatie kan ontbreken. Daarnaast kan informatie ongestructureerd worden weergegeven, niet consistent zijn ten opzichte van andere documentatie en er kan sprake zijn van elite-bias. Onder elite-bias wordt het materiaal bedoeld dat afkomstig is van mondige, goedgeïnformeerde, en doorgaans vooraanstaande respondenten waarbij de bewijsvoering snel als juist wordt gezien (Bergsma, 2003). De onderzoeker heeft kritisch naar de gevonden documentatie gekeken. Per document is bepaald: wat het doel van het document is, de omstandigheden van de totstandkoming van het document, voor wie is het document bedoeld en of er andere informatie is die een beeld kan bevestigen of ontkrachten (Jochems & Joosten, 2005).

"Een interview is een doelgericht gesprek tussen twee of meer personen." (Kahn & Cannel, 1957). Gedurende de meervoudige casestudy zijn personen bevraagd middels semigestructureerde en ongestructureerde interviews. Het interview met personen die niet direct werkzaam zijn bij Business Intelligence is ongestructureerd. Het interview met personen die wel werkzaam zijn bij Business Intelligence is semigestructureerd van aard. De interviews kenmerken zich door een vooraf opgezette topiclijst waarbij gedurende het interview ruimte was om vervolgvragen te stellen, vragen weg te laten, de volgorde van vragen aan te passen en extra vragen te stellen. De interviews binnen de casestudies hebben primair face to face, een op een, plaatsgevonden. Per casus is de richtlijn gehanteerd om minimaal 3 interviews te houden. De beoogde respondenten per casus zijn een manager, een architect, en een gebruiker/verwerker van de Open Data. De manager doet op strategisch niveau zijn verhaal over Open Data. De architect doet op business, informatie en technologisch vlak een invulling. En de gebruiker van Open Data geeft hands-on informatie. Tijdens de interviews heeft de onderzoeker audio opnamen gemaakt. Voor ieder interview is 1 tot 1,5 uur tijd gepland. Voor de meervoudige casestudy zijn drie organisaties gezocht die als respondent willen dienen. Doel is om twee organisaties die Open Data gebruiken te bevragen, en een onderneming te bevragen die géén gebruik maakt van Open Data.

2.2.3. Betrouwbaarheid, validiteit en ethiek

2.2.3.1. Betrouwbaarheid

“Bij het vaststellen van de betrouwbaarheid van onderzoeksresultaten gaat het er om aannemelijk te maken dat een herhaling van het onderzoek tot dezelfde resultaten zal leiden” (Schijns & Oude-de Wolf de, 2004). Het betreft de mate waarin resultaten uit een eerste meting overeenkomen met resultaten uit een tweede meting in het geval er sprake is van dezelfde onderzoekseenheden, hetzelfde meetinstrument en dezelfde context.

Survey

Bij de invulling van de survey kunnen allerlei toevallige fouten optreden, bijvoorbeeld een vraag wordt verkeerd geïnterpreteerd, men vinkt per ongeluk een vinkje verkeerd aan of men veronderstelt dat een ‘Likert’ schaal van laag naar hoog verliep, terwijl de schaal juist van hoog naar laag verliep. Om de betrouwbaarheid van de survey te verhogen is een aantal proefpersonen gevraagd de vragenlijst in te vullen en feedback geven.

Casestudy

De casestudy krijgt vaak kritiek als het gaat om het gebrek aan betrouwbaarheid (Jochems & Joosten, 2005). Om betrouwbaarheid te versterken zijn een tweetal acties uitgevoerd die door Yin (2009) worden voorgesteld. Allereerst is een schriftelijk protocol, zie bijlage D, uitgewerkt voor de uitvoering van de casestudy wat bestaat uit een beschrijving van de documentatiestudie en een beschrijving van de uitvoering van de voorgenomen semigestructureerde en ongestructureerde interviews. Daarnaast is een apart databestand gehandhaafd om de verzamelde documentatie te beheren. Dit databestand heet ‘CommunicatieBeheer’.

2.2.3.2. Validiteit

Validiteit geeft aan in hoeverre gemeten wordt wat gemeten moet worden. Interne validiteit geeft aan in hoeverre het meetinstrument meet wat het verondersteld te meten (Schijns & Oude-de Wolf de, 2004). Externe validiteit wordt vaak ook generaliseerbaarheid genoemd, wat aangeeft in hoeverre resultaten uit een meting veralgemeniseerd worden naar een totale populatie. Het grote verschil tussen interne en externe validiteit betreft dat het respectievelijk om de validiteit gaat van een meetinstrument en de resultaten van een onderzoek.

Survey

Om de interne validiteit te verhogen is aan de afstudeerbegeleider en aan collega-reviewers het verzoek gedaan om te controleren of de vragenlijst van de survey de onderzoeksvragen E1 en E2 voldoende dekt. Dit wordt contentvaliditeit genoemd. De externe validiteit, in mindere mate van belang bij kwalitatief onderzoek, is beperkt. Het gaat immers om een kwalitatieve inventarisatie. Met een steekproefmarge van 5%, een populatie van 500 en een uitkomst van 50% is bij een betrouwbaarheidsniveau van 95% een steekproefgrootte van 217 (43,4%) noodzakelijk. Dit lijkt niet haalbaar. Voor de respons wordt een minimum van 50 van de 500 nagestreefd.

Casestudy

Om de interne validiteit van de casestudy te bevorderen is rekening gehouden met het 'control effect'. Door de aanwezigheid van de onderzoeker kunnen bepaalde verschijnselen worden veroorzaakt. Voorafgaand aan interviews is aan de te interviewen personen gevraagd om, wanneer zij neigen naar bepaalde antwoorden door invloed van de onderzoeker, dit te melden. Ter plekke zou dit dan besproken worden waarna het gesprek wel al dan niet voortgezet zou kunnen worden. Naast dit effect wordt ook gelet op het 'biased viewpoint effect'. Bij beperkte informatie trekt de onderzoeker geen conclusies maar vraagt hij door zodat volledige informatie ontstaat. Selectieve perceptie is hierin tegen gegaan door steeds na een interview, de digitale beschrijving te delen met de geïnterviewde met het verzoek de beschrijving te accepteren 'as-is'. Tenslotte is de interne validiteit verder versterkt door een maatregel die Yin (2009) voorstelt. De onderzoeker heeft getracht patronen te herkennen over de verschillende casestudies heen. Daarbij is de tooling 'NVivo' ondersteunend geweest.

De externe validiteit van de casestudy is versterkt door de uitvoering van een meervoudige casestudy. Daarbij is getracht een differentiatie (branche, grootte) aan te brengen om breder te kunnen generaliseren. Bij de casestudies is replicatie logica (Yin, 2009) toegepast om de externe validiteit te verhogen. Het betreft hier geen statistische maar een analytische generalisering.

Naast interne en externe validiteit wordt ook gesproken over constructvaliditeit. Construct- of begripsvaliditeit betekent dat de interpretatie van de gegevens uit een onderzoek zodanig moet zijn dat kan worden overgegaan van het niveau van empirische verschijnselen naar dat van theoretische begrippen (Jochems & Joosten, 2005).

Om de constructvaliditeit, specifiek voor de casestudy, te verhogen zijn een aantal acties genomen (Yin, 2009):

- Er zijn meerdere bewijsbronnen gebruikt. Enerzijds zijn er meerdere casestudies gedaan, anderzijds hebben er zowel interviews plaatsgevonden als documentatiestudies.
- Interviews zijn digitaal vastgelegd en opgeslagen.
- De respondenten hebben een review uitgevoerd op de digitale vastlegging.

2.2.3.3. Ethiek

Gedurende het onderzoek heeft de onderzoeker met verschillende ethische problemen te maken gehad. Hiermee wordt bedoeld de correctheid van het gedrag van de onderzoeker ten opzichte van de rechten van degenen die het onderwerp zijn van het onderzoek, of de effecten daarvan zullen ondervinden (Saunders et al., 2011).

Er zijn een aantal algemene ethische kwesties waarmee de onderzoeker rekening heeft gehouden:

- Voor zowel de survey als de casestudy geldt dat deelname vrijwillig is en dat men het recht heeft om zich deels of volledig terug te trekken.
- Naar wens kunnen respondenten anoniem blijven.

Survey

Gedurende de uitvoering van de survey blijven de respondenten anoniem. Dat wil zeggen dat namen van organisaties niet vermeld worden. Bij de survey is gezorgd dat niet herleidbaar is welke respondent welke keuzes heeft gemaakt in de survey. Het webscript dat gebruikt wordt voor de survey geeft dit ook specifiek aan. Aan het einde van de survey is gevraagd of de onderneming zich wil aan bieden als kandidaat voor de casestudy. Participatie kon plaatsvinden door zich aan te melden via het emailadres van de onderzoeker. Het emailadres is niet gekoppeld aan de ingevulde survey.

Casestudy

De wens van de onderzoeker is om de naam van de onderneming en die van de respondenten te kunnen vermelden. Op verzoek van de onderneming en de respondenten kan anonimisering toegepast worden. In alle gevallen is toestemming van de deelnemers en de organisatie gevraagd om met naam te mogen publiceren.

Gedurende de casestudy is met verschillende personen gesproken over de toepassing van Open Data door een Business Intelligence afdeling. Non-verbale reacties van respondenten zijn niet schriftelijk.

De onderzoeker heeft zich naar behoren gedragen en zich geconformeerd aan de huisregels van de organisatie waar de casestudies hebben plaatsgevonden. Voorafgaand aan de casestudy is ook verzocht om regels, en daar waar regels niet correct werden nageleefd heeft de onderzoeker aan de respondent gevraagd dit direct aan te geven.

De organisaties die aan het onderzoek meewerken, zijn geen branchegenoten. Dat wil zeggen dat bij de selectie van respondenten geen concurrenten zijn gekozen.

In de documentatiestudie zijn verschillende documenten bestudeerd. Deze documenten worden niet met derden gedeeld worden en zullen vertrouwelijk worden behandeld. Vijf jaar na afronding van het afstudeeronderzoek zullen de documenten die beschikbaar zijn gesteld, worden verwijderd uit de bestandsopslag van de onderzoeker. De onderzoeker is verplicht documentatie gedurende vijf jaar te bewaren.

2.2.4. Wijze van analyseren

Survey

De kwalitatieve gegevens uit de survey zijn na de openstellingsperiode gedownload van de website van de onderzoeker. Er zijn verschillende formaten waarin de gegevens gedownload kunnen worden. Dat betreft formaten als XML, TXT (tab divided) en LSS. Deze formaten kunnen in de meeste analyseprogramma's worden ingelezen als Excel, SPSS, etc. De onderzoeker heeft de data in Excel gedownload.

De gegevensattributen zijn van categorische, nominale, dichotome en ordinale aard. Zij zijn na de download inductief / interpreterend gecodeerd. Analyse van de resultaten is gebeurd in het programma Excel. In Excel zijn eerst alle individuele variabelen onderzocht, getypeerd en met elkaar vergeleken waarna ze gepresenteerd zijn. Daarnaast is gebruik gemaakt van de kwalitatieve onderzoekstool NVivo.

Casestudy

De kwalitatieve analyse bestaat uit een aantal stappen die zijn uitgevoerd in de tooling NVivo. Nadat de interviews zijn getranscribeerd, zijn ze in de tooling geladen. Vervolgens is er een boomstructuur, zie bijlage J, opgezet naar de opbouw van de operationalisering zoals weergegeven in bijlage G. Hierna zijn de interviewteksten achtereenvolgens open en axiaal gecodeerd. Dat wil zeggen dat de teksten uitgesplitst zijn in eenheden, daarmee gekoppeld aan 'nodes' in NVivo waarna eenheden, 'nodes', zijn vergeleken. Daar waar eenheden zichtbaar zijn die niet in de opgezette boomstructuur vooraf waren gedefinieerd, zijn deze alsnog toegevoegd.

De documenten uit de documentatiestudie zijn op gelijke wijze behandeld in NVivo.

2.2.5. Verantwoording survey en casestudy

Voor de vastlegging van communicatie omtrent survey en casestudy is een Excel sheet gebruikt, 'CommunicatieBeheer'. Hierin zijn alle contacten vastgelegd die tijdens het onderzoek hebben plaatsgevonden.

Survey

De aangeschreven organisaties betreffen de 500 grootste private organisaties van Nederland. Met privaat wordt bedoeld commerciële organisaties. Er worden geen publieke/ (semi-)overheid organisaties aangeschreven. De aanschrijving heeft plaatsgevonden op basis van de gepubliceerde lijst in Elsevier, 2004. Aan iedere organisatie is gevraagd om één afgevaardigde van Business Intelligence de survey in te laten vullen. Per organisatie is de naam, activiteit, omzet, winst/verlies en het aantal werknemers beschikbaar.

Casestudy

De organisatie waar de onderzoeker werkzaam is, Achmea, heeft geparticipeerd in de casestudy. Daarbij heeft de afdeling Communicatie aangegeven akkoord te zijn met het publiceren van de naam Achmea. Daarnaast heeft de organisatie Enexis en een anonieme organisatie geparticipeerd in de casestudy.

3. Literatuurstudie

3.1. Literatuur Open Data en Business Intelligence

Via de Digitale Bibliotheek van de Open Universiteit is geprobeerd literatuur te vinden over de toepassing van Open Data door Business Intelligence. De toepassing van Open Data in het algemeen is beperkt beschreven in de literatuur. Over de toepassing van Open Data door Business Intelligence blijkt geen literatuur beschikbaar te zijn. Om vast te stellen dat daadwerkelijk géén literatuur aanwezig is, zijn een aantal stappen genomen ter bevestiging.

- Er is met verschillende partijen contact gezocht met het verzoek om literatuur aan te wijzen. Daarbij zijn benaderd; de aanbieder van Open Data van de Nederlandse overheid, organisaties die diensten aanbieden op basis van Open Data en personen die middels publieke initiatieven het gebruik van Open Data willen bevorderen.
- Er is in de LinkedIn groep 'Open Data Research Network' een bericht geplaatst met het verzoek om literatuur aan te wijzen. Daar zijn zo'n 1000 mensen lid van.
- Er is contact opgenomen met de auteurs van verschillende wetenschappelijke artikelen over Open Data die zijn gebruikt bij de literatuurstudie. Zie bijlage E voor hun reacties.

Deze stappen hebben géén literatuur opgeleverd.

3.2. L1 – Wat is Open Data?

Om deze vraag te beantwoorden, is gekeken naar de losse termen 'Open' en 'Data'. Vervolgens is de definitie van Open Data, zoals gebruikt in het onderzoek, gegeven. Tenslotte volgt een weergave van verschillende gerelateerde concepten en de verschillende categorieën Open Data.

3.2.1. Open Movements

Naast de 'Open Data Movement' zijn verschillende andere 'Open Movements' gaande. Deze 'Open Movements', letterlijk vertaald 'open bewegingen', hebben een gelijk streven als het gaat om openheid. Hieronder worden bekende 'Open Movements' besproken:

- De 'Open Access movement' streeft naar vrije toegankelijkheid als het gaat om wetenschappelijke / academische werken. Het inzien van deze werken neemt vaak kosten met zich mee, sterker nog, het is meestal erg duur (Cirasella, 2014). Open Access streeft er naar dat iedereen in de gelegenheid is om wetenschappelijke werken te raadplegen. Auteurs worden door sommige Open Access journals wel gevraagd om een publicatie-vergoeding te doen (Kozak & Hartley, 2013).
- Open Content betreft de beweging die streeft naar vrije toegankelijkheid van werken met de prominente eigenschap dat werken onder vrije licentie worden gepubliceerd.
- De 'Open Knowledge Movement' staat voor het toegankelijk maken en verspreiden van kennis. Hierbij is kennis een paraplubegrip. Het is een superset van andere 'Open Movements' als Open Source, Open Content, etc. (García - Peñalvo, García de Figuerola, & Merlo, 2010).

- De bekendste ‘movement’ is toch wel die van de ‘Open Source’. Open source software is software waarbij het licentiemodel uitgebreide rechten verleend om de software en de broncode te gebruiken, te modificeren en te herdistribueren zonder de vereiste dat er royalty’s betaald moeten worden (Mijinyawa & Abdulwahab, 2014).

De gedeelde factor van deze ‘Movements’ betreft openheid. De Open Knowledge Foundation (OKFN) is een non-profit netwerk met als doel deze openheid te ondersteunen en te promoten. De OKFN werkt aan verschillende projecten om deze doelstelling te behalen. Een van deze projecten betreft de ‘Open Definition’ waarbij middels de website OpenDefinition (www.opendefinition.org) de samenvattende term voor ‘Open’ wordt gedefinieerd in 37 verschillende talen. Deze term luidt: “Open betekent dat iedereen vrije toegang heeft, vrij kan gebruiken, vrij kan bewerken en vrij kan delen voor ieder doel”. Deze samenvattende term kan verder verbijzonderd worden naar ‘Open Works’¹ en ‘Open Licenties’. Open Works concentreert zich op de toegankelijkheid en een open formaat terwijl Open Licenties zich meer richt op de rechten en plichten die verbonden zijn aan het (her)gebruik.

3.2.2. Data, gegevens en informatie

De computerwetenschap berust op de scheiding van data en software (Tammisto & Lindman, 2012). Waar de ‘Open Source Movement’ zich bezig houdt met de licenties van software, houdt de Open Data beweging zich bezig met de toegang en het hergebruik van data (Lindman, 2014). Een synoniem dat gebruikt kan worden voor data is gegevens. Onder data óf gegevens wordt de vastgelegde uitdrukking van een of meerdere feiten verstaan. De definitie van gegevens wordt vaak over een kam geschoren met de definitie van informatie. Echter is er een wezenlijk verschil, gegevens worden pas informatie als zij door mensen begrepen worden en van nut zijn. Met andere woorden, informatie betreft een zinvol resultaat van de verwerking van gegevens. Volgens Michalewicz (2007) worden data verzameld op dagelijkse basis in de vorm van bits, nummers, symbolen en objecten. Terwijl informatie georganiseerde data is, voorwerkt, geschoond, in structuur gegoten en ontdaan van redundantie.

De data in de term Open Data wordt gezien als gegevens. Open Data wordt veelal aangeboden in twee vormen. In batchvorm waarbij een bestand in een volledige set records wordt aangeboden, en in single-record vorm waarbij vaak gebruik wordt gemaakt van Application Programming Interfaces (API) om data op te halen.

¹ De definitie *Works* wordt gebruikt om een item of deel van kennis aan te duiden.

3.2.3. Definitie van Open Data

Er zijn verscheidene definities van Open Data in de literatuur te vinden. In de kern komen de definities overeen. In tabel 5 een overzicht van de definities:

Definitie	Auteurs
<i>“Open Data is data die vrij gebruikt kan worden, hergebruikt kan worden en opnieuw verspreid kan worden door iedereen – onderworpen enkel, in het uiterste geval, aan de voorwaarde dat men de bron vermeld en de data ook weer deelt. “</i>	(Lindman, 2014) (Ubaldi, 2013) (Borglund & Engvall, 2014) (“The Open Data Handbook” 2014) (“The Open Knowledge Foundation” 2014)
<i>“Open Data zijn vrij toegankelijk online, beschikbaar zonder technische restricties voor hergebruik, en voorzien in een open licentie die toestaat dat data hergebruikt kan worden zonder beperking, binnen verschillende werkvelden.”</i>	(Jetzek, Avital, & Bjorn-Andersen, 2014)
<i>“Open Data zijn vrij toegankelijk online, beschikbaar voor iedereen om te gebruiken en naar wens te opnieuw te publiceren, zonder auteursrechtelijke beperkingen, patenten of andere controle mechanismen.”</i>	(Lindman, 2014)
<i>“Open Data zijn herbruikbaar, vrij toegankelijk, begrijpelijk, toetsbaar.”</i>	(Hester, 2014)

Tabel 5 - Open Data definities

Het overgrote deel van de literatuur gebruikt de eerste definitie als dé definitie voor Open Data. Daarbij wordt deze definitie ook ondersteund door de Open Knowledge Foundation. Binnen het onderzoek zal gebruikt worden gemaakt van deze definitie.

3.2.4. Gerelateerde concepten

Lindman (2014) en Ubaldi (2013) geven aan dat er verschillende concepten zijn gerelateerd aan het concept van Open Data. Welke concepten dat zijn, en hoe ze gerelateerd zijn aan Open Data, wordt in deze paragraaf uitgewerkt. De concepten worden met regelmaat, onterecht, in de literatuur gebruikt als synoniem. Zo zien sommige auteurs bijvoorbeeld Public Sector Information als synoniem voor Open Data (Borglund & Engvall, 2014).

Open Government Data

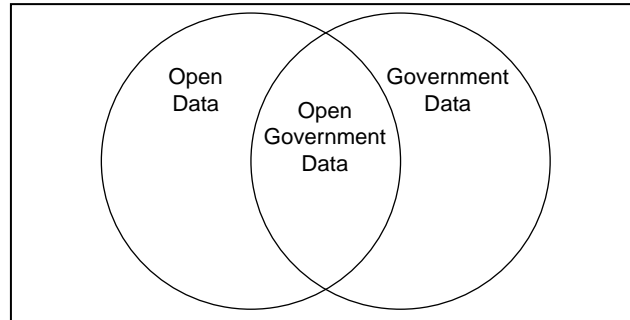
Open Data kunnen door iedereen worden vrijgegeven. Een commerciële organisatie, een overheidsinstantie, zelfs particulieren kunnen datasets openbaar maken (Lindman, 2014).

Open Government Data, zie figuur 4, bevat twee elementen: Government Data en Open Data.

Het element Open Data is eerder omschreven als “Data die vrij gebruikt kan worden,

hergebruikt kan worden en opnieuw verspreid

kan worden door iedereen – onderworpen enkel, in het uiterste geval, aan de eis tot het toeschrijven van de bron en eveneens het gelijk delen er van. “. Government Data betreft alle data en informatie gepubliceerd in opdracht van overheidsinstanties (Ubaldi, 2013). Hieruit kan afgeleid worden dat Open Government Data de Open Data betreft die in opdracht van overheidsinstanties is gepubliceerd.



Public Sector Information

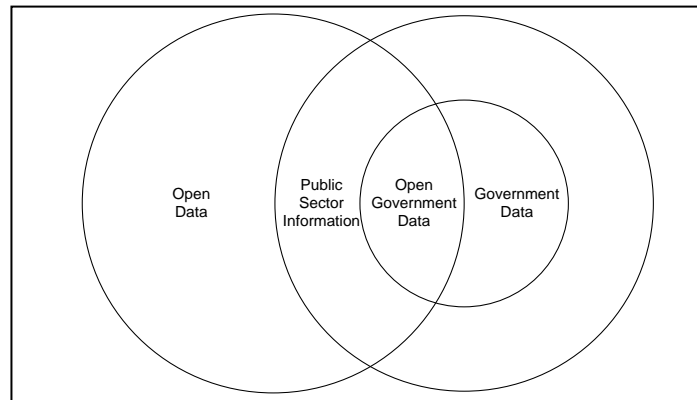
Public Sector Information is breder en omvat

Open Governance Data. Public Sector

Information wordt gedefinieerd als

“Informatie, inclusief informatie producten en diensten, gegenereerd, gemaakt, verzameld, verwerkt, bewaard, onderhouden, verspreid, of gefinancierd door of voor een overheid of publiekelijke instelling” (Ubaldi, 2013).

Daaraan toevoegend zijn er twee sleutel kenmerken bij het hergebruiken van Public Sector Information.



Enerzijds zijn gebruikers individuen of organisaties en anderzijds is het doel van het gebruik om de behoefte van de gebruikers anders dan het originele gebruikers te bevredigen (Hongyan & Mingdi, 2012). Figuur 5 geeft een weergave van public sector information in relatie tot Open Government Data en Open Data.

Big Data

De exacte definitie van Big Data is moeilijk te bepalen. In artikelen worden vaak verschillende termen

gegeven. De term Big Data is ‘vaag’ en wordt vaak gecoöpteerd voor andere doeleinden (Mackie, Sim, &

Johnman, 2015). Ubaldi (2013) geeft aan dat Big Data een verzameling van datasets is, zo groot en complex dat het moeilijk is om voorhanden database management tools te gebruiken, of traditionele data verwerkings applicaties, voor het verwerken van de data met inbegrip van vastlegging, opslag, zoeken, delen, overdracht, analyse en visualisatie. Tenminste drie concepten worden altijd genoemd als het gaat om Big Data. Het grote

data volume, de snelheid waarmee de data verwerkt moet worden en de diversiteit van de data (Pflugfelder, 2013) (Watson, 2014). In het boek 'De Big Data Revolutie' (Mayer-Schönberger & Cukier, 2013) worden drie andere factoren genoemd van Big Data. Ten eerste de factor variëteit waarbij verschillende bronnen aanwezig zijn die elkaar tegenspreken en het gebruik van Big Data compliceren. Een tweede factor is de kwaliteit van data. Kwaliteit van data hoeft niet 100% te zijn. Big Data wordt gekenmerkt door rommeligheid. De derde factor betreft de complexiteit van data. De data is ongestructureerd en afkomstig uit verschillende bronnen. Kijkend naar de begrippen Open Data en Big Data dan bestaat er een overlap, data kan immers big én open zijn. Door literatuur worden Open en Big data regelmatig zij aan zij besproken. Belangrijk is dat het geen synoniemen zijn.

3.2.5. Categorieën Open Data

Het project 'Open Data Barometer' richt zich op het achterhalen van de impact van de wereldwijde Open Data initiatieven. Open Data betreft data die over uiteenlopende onderwerpen gaat met verschillende potentiële mogelijkheden. De Barometer presenteert de meest voorkomende categorieën van Open Data. De categorieën zijn verdeeld over drie hoofdcategorieën, te weten (Davies, 2013):

1. Innovatie: Open Data die gebruikt kunnen worden in nieuwe organisaties, of van waarde is binnen bestaande organisaties. Subcategorieën:
 - a. Kaartgegevens
 - b. Dienstregelingen openbaar vervoer
 - c. Misdaadcijfers
 - d. Internationale handelsgegevens
2. Sociaal beleid: Open Data die nuttig is in relatie tot planning, levering, bekritisering en verbetering van sociaal beleid.
 - a. Prestaties van de gezondheidssector
 - b. Prestaties van lager of secundair onderwijs
 - c. Nationale milieu statistieken
 - d. Volkstellingsgegevens
 - e. Gegevens over grondbezit
3. Verantwoording: Open Data die er voor zorgt dat overheden en organisaties verantwoordelijk gesteld kunnen worden.
 - a. Wetgeving
 - b. Verkiezingsresultaten
 - c. Regeringsbudget
 - d. Regeringsuitgaven
 - e. Bedrijvenregister

Het is belangrijk te melden dat de verschillende subcategorieën niet per definitie aan de hoofdcategorie toebehoren. Er bestaat een zekere overlap. Zo kunnen volkstellingsgegevens bijvoorbeeld ook ter innovatie benut worden. Het is niet zalmakend, een richtlijn.

3.3. L2 – Wie stelt Open Data beschikbaar?

3.3.1. Bronnen

Open Data kunnen door iedere persoon en door iedere (private/publieke) organisatie beschikbaar worden gesteld. Er zijn twee grote bronnen zichtbaar als het gaat om het aanbieden van Open Data.

Open Data uit wetenschappelijke bron wordt Open Science Data genoemd. Open Science Data richt zich op de publicatie van data en resultaten van wetenschappelijk onderzoek. Een van de grootste belemmeringen voor de publicatie van Open Science Data betreft de wens van onderzoekers om data te publiceren aan de gemeenschap maar de beperking daarop vanuit de uitgevers die rechten over de data claimen en het niet toestaan om de data te hergebruiken (Murray-Rust, 2008).

Toch is er een publicerende bron zichtbaar die de overgrote hoeveelheid Open Data beschikbaar stelt en de overige bronnen in de schaduw plaatst. Dat betreft de Open Data, gepubliceerd door overheidsinstanties. In de context van Open Data publicaties vanuit de overheid worden vaak de begrippen Open Government Data en Public Sector Information gebruikt als synoniem voor Open Data. Zo bijvoorbeeld ook de Open Data website (data.overheid.nl) van de Nederlandse overheid die Open Data definieert als “data bekostigd uit publieke middelen”, waar eigenlijk de term Open Government Data beter past. Opvallend is ook dat het voorzien in Open Data vooral een nationale, en niet een lokale, aangelegenheid is. Ruim 77 landen publiceren momenteel Open Data (Davies, 2013).

3.3.2. Redenen van openstelling

Overheidsinstanties publiceren Open Data om verschillende redenen (Davies, 2013).

1. Het efficiënter en effectiever maken van overheden. Zowel door de overheid die zijn eigen data gebruikt om beter te gaan werken, als externe partijen die overheden helpen te versnellen en verbeterde publieke diensten te leveren.
2. Stimuleren van innovatie en economische groei. Zowel startende organisaties als reeds bestaande organisaties kunnen hun voordeel doen met de Open Data die door overheden wordt gepubliceerd. Zij kunnen nieuwe en/of vernieuwde producten en diensten bieden, om omzet te verkrijgen/te verhogen.
3. Transparantie en verantwoording. Dit zorgt er voor dat mensen het handelen van overheidsinstanties kunnen monitoren. Daarbij vergemakkelijkt het publiek om overheidsinstanties te controleren op hun werkzaamheden, eventuele corruptie en andere onverklaarbare activiteiten.

Huijboom & Van den Broek (2011) omschrijven soortgelijke redenen van openstelling:

1. De toename van democratische controle en politieke deelneming. Burgers kunnen beter hun democratische rechten uitoefenen en kunnen verder betrokken raken bij politiek en beleidsvoering.
2. Bevorderen van product en dienst innovatie, waar samenhang is met punt 2 van Davies (2013). ICT organisaties zijn in staat om nieuwe diensten te creëren en burgers kunnen hun ideeën en creativiteit in praktische oplossingen omzetten.
3. Versterking van de wetshandhaving. Wat zich richt op het versterken van de veiligheid binnen een land.

3.3.3. Promotie gebruik Open Data

Open Data wordt gepubliceerd op het internet, maar hoe wordt Open Data aan de man gebracht? Er zijn verschillende instrumenten die het mogelijk maken om de afname van Open Data te bevorderen.

Open Data Ecosystemen

Ubaldi (2013) geeft aan dat het belangrijk is een Open Data Ecosysteem te creëren dat het mogelijk maakt om verschillende Open Data te verkennen. De definitie van een Open Data ecosysteem is “een multi-niveau en multidimensionale entiteit waar ruwe materialen, voor wat betreft de distributie en ontwikkeling van belang zijn, het doel is van samenwerking (Zuiderwijk, Janssen, & Davis, 2014)”. Kort gezegd een systeem of voorziening die Open Data aanbiedt.

Open Data Ecosystemen hebben een grote hoeveelheid karakteristieken en zijn door verscheidene auteurs besproken. De karakteristieken zijn te categoriseren naar ontwerp, context, onderlinge afhankelijkheden & interactie, deelnemerschap en databronnen & tools (Zuiderwijk et al., 2014). Zuiderwijk geeft tevens aan dat een Open Data ecosysteem aan vier basale elementen zou moeten voldoen.

1. Vrijgave en publicatie van Open Data op het internet.
2. Zoeken, vinden, evalueren en bekijken van data en de gerelateerde licenties.
3. Schonen, analyseren, verrijken, combineren, linken en visualiseren van data.
4. Begrijpen en bediscussiëren van data, en het voorzien in feedback naar de data leverancier en andere stakeholders.

Ter integratie van de elementen van een ecosysteem oppert Zuiderwijk (2014) nog een drietal elementen toe te voegen.

5. Het opleiden van gebruikers in de manier waarop Open Data gebruikt kan worden.
6. De implementatie van een kwaliteit management systeem met als doel om de kwaliteit van data te beoordelen.
7. Verschillende typen metadata ter bevordering van de interoperabiliteit en verwerking.

Instrumentarium

Huijboom en Van den Broek (2011) geven een aantal andere instrumenten om Open Data strategieën te bevorderen. Deze instrumenten zijn gecategoriseerd naar:

1. Educatie en training. Het aansporen van overheidsinstanties om Open Data te publiceren door ze hierover in te lichten en te trainen.
2. Vrijwillige benaderingen. Op vrijwillige basis voorzien in adviezen en publiekelijke vrijwilligheid.
3. Economische instrumenten. Het houden van competities, hackactons, unconferences, camps, etc. Anderzijds het financieren van Open Data portalen.
4. Wetgeving en controle. Er voor zorgdragen dat er een open wetgeving wordt gehanteerd bij vrijgave van Open Data. Het nastreven van open technische standaarden die interoperabiliteit stimuleren. Het monitoren van de voortgang van Open Data initiatieven.

3.4. L3 – Wat zijn de kenmerken van Open Data?

3.4.1. Kenmerken van Open Data

Vanuit verschillende conceptperspectieven worden de kenmerken van Open Data toegelicht. Hieronder volgt een overzicht van de kenmerken van Open Data.

Lindman (2014) stelt dat er een drietal categorieën van openheid zijn in de context van Open Data.

1. Technische openheid: Open Data wordt beschikbaar gesteld met een juiste interface beschrijving en volgens de juiste standaard.
2. Wettelijke openheid: Open Data wordt zonder belemmeringen vrijgesteld om herbruikt te worden.
3. Commerciële openheid: Open Data is vrijgesteld voor commercieel gebruik.

Borglund & Engvall (2014), Ubaldi (2013) en Bauer & Kaltenböck (2011) geven de volgende kenmerken van Open Data.

1. Compleet: Open Data wordt compleet geleverd.
2. Primair: Open Data worden bij de bron verzameld met het hoogste mogelijk niveau van granulariteit.
3. Tijdig: Open Data worden zo spoedig mogelijk beschikbaar gemaakt om de waarde van de data te behouden.
4. Toegankelijk: Open Data zijn beschikbaar voor iedere geïnteresseerde en voor elk denkbaar doel.
5. Machine verwerkbaar: Open Data zijn gestructureerd om automatische verwerking te bespoedigen.
6. Niet discriminerend: Open Data zijn voor iedereen beschikbaar zonder vooraf te hoeven registreren. Specifieke personen en/of groepen worden niet uitgesloten om welke reden dan ook.
7. Niet gebonden aan leveranciers: Data zijn beschikbaar in een formaat waar geen entiteit de exclusieve controle over heeft.
8. Licentievrij: Open Data hebben geen auteursrechten, patenten, copyrights, etc.

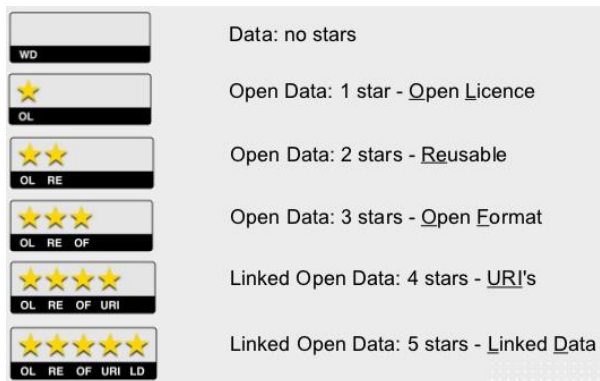
De kenmerken van Lindman (2014) kunnen afgezet worden ten opzichte van de kenmerken van Borglund & Engvall (2014) en Ubaldi (2013). Zie afbeelding 6.

	Compleet	Primair	Tijdig	Toegankelijk	Machine verwerkbaar	Niet discriminerend	Niet gebonden	Licentie vrij
Technische openheid	X	X	X	X	X			
Wettelijke openheid								X
Commerciële openheid						X	X	

Tabel 6 - Kenmerken Lindman versus kenmerken Borglund&Engvall en Ubaldi

3.4.2. Mate van hergebruik

Tim Berners-Lee wordt door velen gezien als de grondlegger van het World Wide Web. Hij richtte het World Wide Web Consortium (W3C) op dat de webstandaarden ontwerpt zoals HTML, XML en CSS. Het algemene doel van Open Data is niet om het alleen maar voor iedereen toegankelijk te maken, maar ook om het herbruikbaar te maken (Borglund & Engvall, 2014). Om deze herbruikbaarheid van Open Data te kunnen kenmerken, ontwikkelde Tim Berners-Lee het vijfsterrenmodel zoals weergegeven in figuur 6 (Erkimbaev, Zitserman, Kobzev, Serebrjakov, & Teymurazov, 2014).



Het vijfsterrenmodel geeft een weergave van de mate van hergebruik. Een van de meest beschreven kenmerken van Open Data, wat toebehoort aan de eerder genoemde categorie technische openheid.

- Een ster: Betreft het beschikbaar maken van data op het web. Ongeacht het formaat en de (on)gestructureerdheid, maar wel onder een open licentie. Voorbeeld: een PDF met het jaarverslag van een organisatie.
- Twee sterren: Betreft het beschikbaar maken van data op het web in een gestructureerd formaat waarbij de data is gebonden aan eigendomsrechten. Voorbeeld: een Microsoft Excelsheet met orderregels.
- Drie sterren: Betreft het beschikbaar maken van data op het web in een gestructureerd formaat waarbij de data niet is gebonden aan eigendomsrechten. Voorbeeld: een CSV ²bestand met orderregels.
- Vier sterren: Betreft het beschikbaar maken van data op het web in een gestructureerd formaat waarbij de data niet is gebonden aan eigendomsrechten en waarbij unieke URI's gebruikt worden om te kunnen verwijzen naar de dataset. Een URI is een Uniform Resource Identifier, een internet-protocolelement dat formeel beschrijft hoe de bron benaderd kan worden. Voorbeeld: een hyperlink naar een CSV bestand met orderregels.
- Vijf sterren: Betreft het beschikbaar maken van data op het web in een gestructureerd formaat waarbij de data niet is gebonden aan eigendomsrechten, waarbij unieke URI's gebruikt worden om te kunnen verwijzen naar de dataset en waarbij de data gelinkt wordt naar andere data om te voorzien in context. Vijf sterren Open Data wordt ook wel Linked Open Data (LOD) genoemd.

3.4.3. Linked Open Data

Het volledige potentieel van Open Data wordt behaald als Open Data overgaat in Linked Open Data (LOD), wanneer Open Data vijf sterren scoort in het vijfsterrenmodel van Tim Berners-Lee. LOD is gebaseerd op een

² CSV staat voor Comma Separated Values. Een CSV-bestand wordt gebruikt voor tabelbestanden. Het bestand bestaat uit kommagescheiden tekstgegevens welke gemakkelijk gebruikt kan worden.

drietal technologieën: HTTP, URI en RDF. Waar HTTP een protocol is voor communicatie tussen webclient en webserver wordt het Resource Description Framework (RDF) gebruikt om uitspraken te doen over kenmerken van bronnen. Dat gebeurt volgens de subject-predicaat-object structuur. Het voordeel van LOD is dat het interoperabiliteit ondersteund van heterogene informatie bronnen (Borglund & Engvall, 2014). Op deze manier kan LOD een bijdrage leveren aan het semantisch web³. Het stimuleert het hergebruik van Open Data optimaal.

In een groot wereldwijd onderzoek waar meer dan 70 landen participeerden, bleek echter dat slechts 9,5% van alle Open (Government) Datasets, Linked Open Datasets te zijn (Davies, 2013). Ook voor de Nederlandse overheid geldt dat het aantal Linked Open Datasets erg beperkt is, zo blijkt uit cijfers van maart 2015. Van de 2844 Open (Government) Datasets zijn geen sets aanwezig die de vijfsterren classificatie, Linked Open Data, van Tim Berners-Lee hebben. Slechts 0,63% van de datasets wordt gekenmerkt met vier sterren, Open Datasets aanwijsbaar middels URI's. Aan kop staan de Open Datasets in open formaat (2784), drie sterren, met 97,8% (data.overheid.nl).

3.4.4. Licentievrij

Een van de kenmerken van Open Data is dat licentievrijheid nagestreefd moet worden, waardoor ze voor iedereen toegankelijk zijn. Voor de openstelling van Open Datasets kunnen verschillende licentiestructuren worden toegepast. Er worden verschillende licentiestructuur aanbevolen bij de publicatie van Open Data. Twee belangrijke typeringingen worden gebruikt bij de opzet van de licentiestructuren, BY en SA. BY, 'requires attribution', staat voor 'vereist verwijzing naar bron' en SA, 'Share-Alike' staat voor 'vereist dat gebruikte Open Data ook gedeeld wordt'. Dit is tevens een rechtstreekse verwijzing naar de definitie van Open Data die aangeeft dat ten hoogste de eis geldt dat er een toeschrijving aan de bron moet plaatsvinden, en dat gebruikte data eveneens gedeeld worden. In tabel 7 staan de aanbevolen Open Data licenties opgesomd ("Open Definition," n.d.).

Licentie	BY	SA
Creative Commons CCZero (CC0)	X	X
Open Data Commons Public Domain Dedication and License (PDDL)	X	X
Creative Commons Attribution 4.0 (CC-BY-4.0)	V	X
Open Data Commons Attribution License (ODC-BY)	V	X
Creative Commons Attribution Share-Alike (CC-BY-SA-4.0)	V	V
Open Data Commons Open Database License(ODbL)	V	V

Tabel 7 - Lijst met soorten licenties bij toepassing Open Data

³ Het semantisch web verschaft een framework waarmee data gedeeld en hergebruikt kunnen worden. Het initiatief is gestart vanuit het World Wide Web Consortium (W3C).

3.5. L4 – Hoe kan Open Data worden toegepast?

Overheden implementeren Open Data strategieën om transparantie, participatie en overheidsefficiëntie te verhogen. Het algemeen aanvaarde uitgangspunt van deze strategieën is dat het publiceren van Open Data vanuit overheden de betrokkenheid van burgers versterkt en tevens zorgt voor nieuwe innovatieve bedrijvigheid (Huijboom & Broek, 2011). Op welke manier kan Open Data innovatief ingezet worden?

Opvallend tijdens het literatuuronderzoek is dat er ontzettend veel literatuur is over de aanbodzijde van Open Data. Echter de wijze waarop Open Data wordt toegepast om meerwaarde te creëren is zeer beperkt. De focus ligt voornamelijk op innovatieve Apps en op eventueel mogelijke bedrijfsmodellen op basis van Open Data.

3.5.1. Websites & Apps

Sinds enkele jaren zijn mobiele apps en web-apps sterk in opkomst. Wereldwijd worden apps beschouwd als zeer belangrijk wanneer het gaat om interactie tussen mensen en bedrijven onderling. Wanneer Open Data gebruikt worden, worden ze vaak ontsloten in websites of apps (Sandoval-Almazan, 2012). De toepassing van Open Data is vooral zichtbaar in de gemeenschapsdienstverlening (Thoreson, 2013).

Er zijn vele online communities van ontwikkelaars die werken met Open Data, en ze onderbrengen in applicaties en apps. Er worden vele bijeenkomsten, hackathons, competities, unconferences en challenges georganiseerd.

3.5.2. Bedrijfsmodel

Vanuit het perspectief van het bedrijfsmodel zijn er drie verdienmodellen met betrekking tot Open Data (Bonina, 2013):

1. Het freemium model: betreft het aanbieden van een gratis basisdienst. Voor extra's moet men betalen.
2. Het kruissubsidiëringsmodel: bij gebruik van dit model vindt prijsdiscriminatie plaats. Een groep gebruikers betaald een hoge prijs voor de dienst om de dienst te subsidiëren en bij een andere groep gebruikers tegen een lagere prijs aan te bieden.
3. Het netwerkeffect model: dit model doelt op samenwerking met andere organisaties. Het doel daarvan is om enerzijds de kosten van de diensten te verlagen en anderzijds de een groter bereik te creëren als het gaat om afnemers van de diensten.

Volgens Ubaldi (2013) zijn er verschillende algemene voordelen te behalen aan het gebruik van Open (Government) Data voor overheid, burgers, de burgermaatschappij en de economie.

Er worden een vijftal business model oerbeelden gegeven voor het creëren van waarde (Ubaldi, 2013):

1. Leveranciers: Zij leveren Open Data.
2. Samenstellers: Zij verzamelen Open Data en combineren deze data om bruikbare inzichten te creëren voor de gebruikers.
3. Ontwikkelaars: Zorgen voor applicaties en diensten met Open Data als bron.

4. Verrijkers: Zijn grote organisaties die grote hoeveelheden Open data gebruiken en combineren met de eigen (gesloten) data om diensten te creëren.
5. Activators: Non-profit organisaties die voorzien in platformen en technologieën voor gebruikers.

3.6. L5 – Wat is Business Intelligence?

De term Business Intelligence bestaat al geruime tijd. Al in het jaar 1865 werd de term bedacht door ene Richard Miller Devens. Hij gebruikte de term om te beschrijven hoe de bankier, Sir Henry Furnese, winst behaalde door het ontvangen en reageren op informatie over zijn omgeving, voordat concurrenten de informatie ontvingen. In het jaar 1958 schreef de onderzoeker Hans Peter Luhn van IBM een artikel waarin hij de term Business Intelligence gebruikte (Luhn, 1958). Hij wordt gezien als grondlegger van de definitie.

Er zijn vele definities van Business Intelligence beschikbaar. Onderzoeksbureau Gartner definieert Business Intelligence als een paraplu-term waarbinnen infrastructuur, software en best practices vallen die het mogelijk maken om toegang te verschaffen tot de analyse van informatie met als doel aan te zetten tot het verbeteren en optimaliseren van besluitvorming en prestatie. Rud (2009) definieert Business Intelligence als **“een verzameling technieken en hulpmiddelen voor de omzetting van ruwe data naar betekenisvolle en bruikbare informatie voor business analyse mogelijkheden. Business Intelligence technologieën zijn in staat om grote hoeveelheden data te verwerken ter identificatie, ter ontwikkeling en ter ontwerp van nieuwe strategische kansen”**.

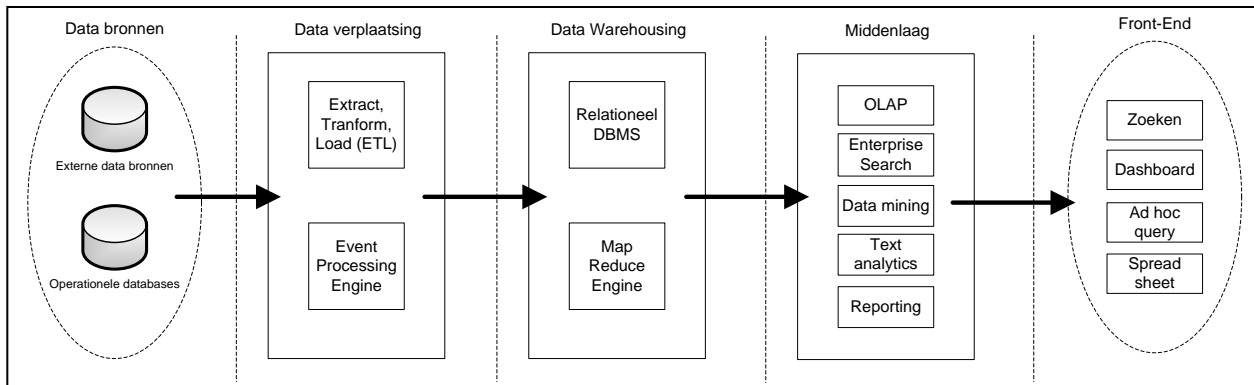
Business Intelligence Software is een verzameling besluitvorming ondersteunende technologieën voor organisaties gericht op het mogelijk maken van kenniswerken waardoor men betere en snellere beslissingen kan maken (Chaudhuri et al., 2011).

Business Intelligence wordt op strategische wijze ingezet. De belangrijke aandachtsgebieden zijn (Negash, 2004)

1. Prestatiemanagement (Corporate Performance Management). Wat zorgt voor inzichten in prestaties van de organisatie, de verschillende afdelingen, de verschillende teams.
2. Het optimaliseren van de relatie met klanten, het monitoren van bedrijfsactiviteiten en ondersteuning in het maken van beslissingen.
3. Standalone Business Intelligence applicaties voor specifieke doeleinden of strategieën.
4. Management rapportages. Rapportages die het management ondersteunen om de organisatie te sturen.

3.7. L6 – Hoe ziet een Business Intelligence architectuur er uit?

Een typische Business Intelligence architectuur staat weergegeven in figuur 7 (Chaudhuri et al., 2011). Deze figuur toont de input, verwerking en output van de ‘klassieke’ Business Intelligence. Interne en externe databronnen worden ingelezen, in een bepaalde vorm opgeslagen in een data warehouse, waarna de gebruiker via verschillende front-end applicaties/diensten een beroep doet op de middenlaag voor de ontsluiting van de gewenste gegevens.



3.7.1. Databronnen en verplaatsing van data

Aan het begin van de Business Intelligence-keten zijn de verschillende interne en externe *databronnen* zichtbaar. De interne bronnen kunnen verschillende operationele databases zijn binnen verschillende afdelingen van een organisatie. Voorbeelden van interne bronnen zijn ERP-, CRM-, SCM-, Marketing-systemen. Voorbeelden van externe bronnen zijn: Data van leveranciers, Open Data, etc. De data uit de diverse bronnen is van gevarieerde kwaliteit en gebruikt niet altijd dezelfde interfacestructuur, dezelfde codetabellen en formaten. Ter gebruik moet de data uit de bron ook worden opgeslagen. Hiervoor zijn ETL-tools beschikbaar die de data uit de bron halen, corrigeren naar de gewenste vorm, en vervolgens laden in een data warehouse. (Chaudhuri et al., 2011). Steeds vaker worden ook Event Processing Engines toegepast om BI taken steeds sneller te kunnen uitvoeren. Dat komt omdat men steeds vaker en sneller wil beschikken over de allerlaatste en meest actuele informatie.

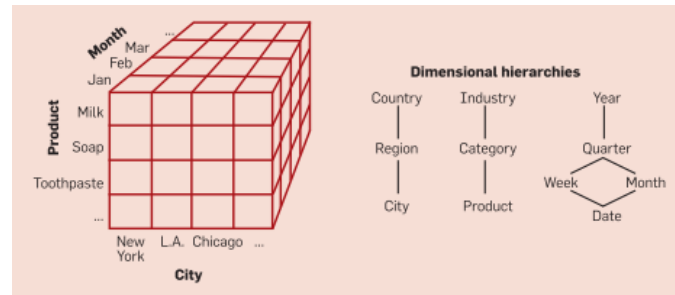
3.7.2. Data warehousing

Nadat data uit de databronnen is geëxtraheerd en bewerkt, moet de data opgeslagen worden. De opslag gebeurt in een zogenaamd data warehouse. Een data warehouse is een centrale opslagplaats voor grote hoeveelheden geïntegreerde data. De meest bekende engine voor het opslaan van data betreft het traditionele ‘Relational Database Management System’, afgekort met RDBMS. Momenteel is de trend dat deze RDBMS'en in-memory en kolom-georiënteerd gaan werken. Te wijden aan de zeer grote hoeveelheden data die tegenwoordig te verwerken zijn. Denk bij deze nieuwe technologie aan SAP HANA (Fischer, 2014). SAP HANA kent zeer hoge transactiesnelheden die tot 1000x hoger liggen dan de traditionele RDBMS'en.

Wanneer grote hoeveelheden data worden verwerkt is het tevens mogelijk om het MapReduce framework te hanteren. Hierbij worden een groot aantal computers gebruikt om parallel data te verwerken.

3.7.3. De middenlaag

Waar de databronnen en de verplaatsing naar data warehousing wordt gezien als een backend proces, bestaat er ook een middenlaag naar de front-end processen. In deze middenlaag bestaan verschillende engines en servers. Een bekende server betreft de OLAP-server. OLAP, staat voor OnLine Analytic Processing. Het biedt de gebruiker multidimensionale views van data waarmee de gebruiker gemakkelijk een selectie van gegevens kan doen vanuit verschillende perspectieven, zie figuur 6. De Enterprise Search engine verzorgt de mogelijkheid om te zoeken binnen de ongestructureerde en gestructureerde data in het data warehouse. Reporting servers maken het mogelijk om rapportages te maken met de data uit het data warehouse als input. Datamining engines maken het mogelijk om data te analyseren en predicatieve modellen te bouwen. Textmining engines kunnen grote hoeveelheden tekst analyseren en waardevolle informatie vinden (Chaudhuri et al., 2011).



Figuur 8 - Klassieke middenlaag Business Intelligence

3.7.4. Front-end toepassingen

Er zijn verschillende soorten toepassingen die werken op basis van de verschillende servers en engines in de middenlaag van de BI-architectuur. Zo wordt het zoeken mogelijk gemaakt, het genereren van spreadsheets, de publicatie van dashboards, ad-hoc te query-en en rapportages te maken.

3.8. L7 – Welke functies kent Business Intelligence?

Business Intelligence kent verschillende functies. De meest gebruikte functies zijn: reporting, OnLine Analytical Processing (OLAP), Analytics, Datamining, Complex Event Processing, Business Performance Management en Textmining (Cebotarean, 2011). Een beschrijving in tabel 8.

Functie	Omschrijving	Auteur
Reporting	Reporting zorgt voor het verzamelen en presenteren van data bij de eindgebruikers. Op basis van de producten van reporting, zoals bijvoorbeeld rapportages, Excel sheets, grafieken en dashboards, worden beslissingen genomen.	(Chaudhuri et al., 2011)
OnLine Analytical Processing (OLAP)	Zoals in paragraaf 4.6.3 aangegeven staat OLAP voor OnLine Analytical Processing. OLAP biedt de gebruiker een gemakkelijke manier om een selectie van gegevens te doen vanuit verschillende perspectieven.	(Chaudhuri et al., 2011)
Analytics	Analytics betreft het ontdekken van betekenisvolle patronen in data.	(Kohavi, Rothleder, & Simoudis, 2002)
Data mining	Betreft het rekenkundige proces van het ontdekken van patronen in grote datasets waarbij verschillende methoden worden gebruikt als artificiële intelligentie, machinaal leren, statistieken en database systemen. Het doel van het data mining proces is om informatie uit een dataset te extraheren en het om te zetten naar een begrijpbare structuur. Data mining betreft de analyse stap in Knowledge Discovery in Databases.	(Chakrabarti, Ester, Fayyad, & Gehrke, 2004)
Complex event processing	Is een verzamelnaam voor methoden, technieken en tools om gebeurtenissen (events) te verwerken zodra ze plaatsvinden. Dat gebeurt continu, tijdig en over verschillende bronnen heen. Hieruit kunnen patronen worden afgeleid.	(Eckert & Bry, 2009)
Business Performance Management	Stelt organisaties in staat om de implementatie van strategische initiatieven effectief te monitoren, beheersen en te managen. Het helpt organisaties om prestaties te optimaliseren.	(Frolick & Ariyachandra, 2006)
Textmining	Textmining is het gebruiken van geautomatiseerde methoden voor het extraheren van informatieve patronen die beschikbaar zijn in teksten.	(Cohen & Hunter, 2008)
Predictive analytics	Men probeert op basis van actuele en historische data de toekomst te voorspellen door gebruik te maken van verschillende statistische technieken. Predictieve modellen maken gebruik van patronen uit historische data om risico's en kansen te identificeren.	(Coker, 2014)

Tabel 8 - Overzicht Business Intelligence functies

3.9. Eindresultaat literatuurstudie

Door uitvoering van de literatuurstudie is getracht vorm te geven aan de manier waarop Open Data binnen een Business Intelligence omgeving toegepast kan worden. Vanuit het Open Data perspectief is verduidelijkt wat Open Data is en wie verantwoordelijk zijn voor het aanbieden er van. Daarbij is ook helder welke eigenschappen en kenmerken Open Data heeft. De manier waarop Open Data worden toegepast wordt beperkt beschreven, dat betreft vooral de toepassing van Open Data in innovatieve, op zich staande, Apps.

Vanuit het Business Intelligence perspectief is verduidelijkt wat de definitie is. Daarbij is verder verduidelijkt hoe een typische Business Intelligence architectuur er uitziet, en welke verschillende functies leven binnen de betreffende architectuur.

Gedurende de literatuurstudie is géén literatuur gevonden over de toepassing van Open Data binnen Business Intelligence. Onderzoek naar de toepassing van Open Data binnen een Business Intelligence organisatie zal deze leemte in kennis vullen als boven water komt hoe de toepassing ervan plaatsvindt.

Het referentiemodel dat opgesteld is naar aanleiding van de literatuurstudie is tweeledig. Het eerste overzicht bevat de elementen van Open Data en het tweede overzicht bevat de elementen van Business Intelligence met daarbij een verwijzing naar Open Data.

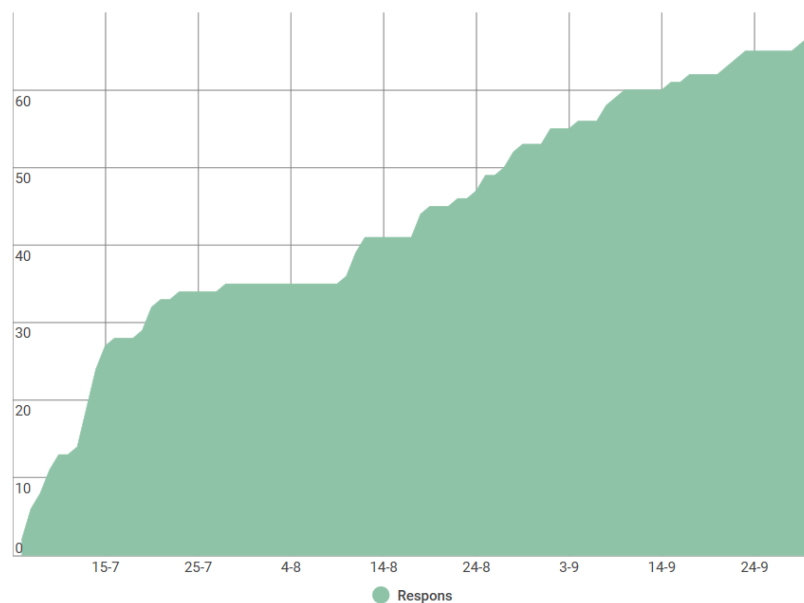
In bijlage E is het referentiemodel uit de literatuur opgenomen, een onderdeel Business Intelligence en ene onderdeel Open Data. Daarnaast is in bijlage F de operationalisering opgenomen.

4. Onderzoeksresultaten

4.1. Survey – Beantwoording E1 en E2

Om antwoord te kunnen geven op de empirische vragen E1 en E2 is een survey ingezet. Uit onderzoek blijkt dat organisaties die Business Intelligence inzetten vaak grote multinationals zijn met verschillende dochterondernemingen (Computerworld, 2006). Daarom betreft de scope van het afstudeeronderzoek de top 500 private organisaties van Nederland. Elsevier heeft in 2014 de namen van de 500 grootste organisaties gepubliceerd. Daarbij zijn verschillende gegevens gepubliceerd als de activiteit van de organisatie, de omzet/winst/verlies cijfers, het aantal medewerkers, etc. De lijst bestaat uit organisaties met een jaaromzet variërend tussen 231 miljoen en 353 miljard en het aantal medewerkers variërend tussen de 110 en 568.000. Voor de survey zijn de organisaties uit deze top 500 lijst benaderd. Het gebruikte survey navigatiepad is te vinden in bijlage B. Bijlage C geeft een overzicht van de gebruikte coderingstabellen en bijlage H geeft een overzicht van de surveyresultaten

In de periode van 29 juni t/m 15 augustus 2015 zijn de top 500 private organisaties van Nederland aangeschreven. Van 29 juni t/m 30 september 2015 zijn de respondenten in de gelegenheid geweest om de survey in te vullen. De respons is in kaart gebracht in figuur 9.



Figuur 9 - Respons in tijd

4.1.1. Respons

De onderzoeker heeft getracht de top 500 private organisaties van Nederland te benaderen. Van de 500 organisaties in scope zijn 127 organisaties, 25,4%, uitgevallen in de aanschrijving. De redenen voor uitval variëren:

- Er was geen contact mogelijk via LinkedIn. Er werden geen personen aangetroffen via de zoekmachine Google of via LinkedIn gelieerd aan de BI afdeling/discipline van de specifieke organisatie.
- Er was geen contact mogelijk via de corporate website. Enerzijds omdat er geen contactgegevens of webformulieren werden aangetroffen, anderzijds omdat de contactgegevens onjuist bleken of dat de webformulieren niet werkzaam bleken.
- Er was alleen contact mogelijk per post of telefoon. Bij sommige organisaties was het niet mogelijk om via de digitale weg contact op te nemen, alleen via de post of per telefoon. Deze contactwijze was uitgesloten.
- De organisatie blijkt te detacheren in IT. Deze organisaties hebben mensen in dienst voor wat betreft de Business Intelligence discipline, maar deze mensen zijn niet op de hoogte van de exploitatie van de interne Business Intelligence.
- Daar waar op de corporate website werd aangegeven dat men geen onderzoeksvragen ter behandeling aanneemt, is de poging tot benadering gestaakt.

Van de 500 organisaties in scope zijn 373 organisaties benaderd. 161 (43,16%) organisatievertegenwoordigers zijn benaderd via LinkedIn. 212 (56,84%) organisaties zijn benaderd via de corporate website. De survey is door 67 organisatievertegenwoordigers ingevuld wat neerkomt op een percentage van 17,69%. De respons op basis van de totale scope is 13,4%.

4.1.2. Achtergrond respondenten

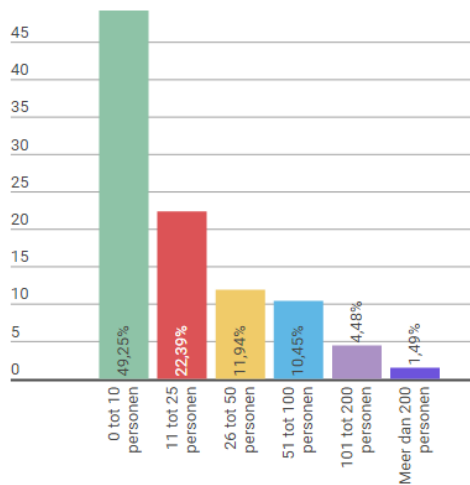
Om een beeld te krijgen van het type respondenten in de survey zijn een tweetal vragen gesteld. Enerzijds de vraag welke functie de respondent uitvoert binnen de Business Intelligence organisatie, anderzijds de dienstjaren van de respondent bij de huidige organisatie.

In 53,7% van de gevallen heeft de respondent de functie van manager/teamleider, 19,4% van de respondenten is consultant en 10,5% is architect. Daarnaast werden in mindere mate andere functies vertegenwoordigd als bedrijfsanalist, informatiemanager, business process analist en technisch/functioneel beheerder. Zie bijlage H.1. voor een genuanceerd inzicht.

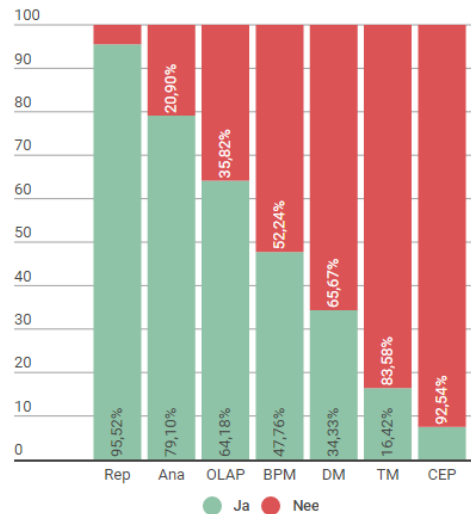
Van de respondenten blijken de meesten al geruime tijd, tussen de 5 en 10 jaar, werkzaam te zijn bij de huidige organisatie, 23,9%. Gevolgd door dienstverbanden tussen de 2 en 3 jaar (20,9%), 0 en 1 jaar (17,9%), 1 en 2 jaar (16,4%) en 3 en 4 jaar (13,4%).

4.1.3. Achtergrond Business Intelligence

Om een beeld te krijgen van het type Business Intelligence afdelingen binnen de organisatie, zijn een vijftal vragen gesteld in de survey. Ten eerste is gevraagd hoe lang de Business Intelligence afdelingen bestaat. De meeste Business Intelligence afdelingen bestaan tussen de 5 en 10 jaar (37,3%). Gevolgd door periodes van 3 tot 4 jaar (16,4%) en langer dan 10 jaar (13,4%). Daarbij zijn organisaties in verscheidene branches actief. De top vertegenwoordigers zijn financiële instellingen (16,4%), groot- en detailhandel (10,5%), Industrie (9%) en vervoer en opslag (9%). In figuur 10 volgt een weergave van de grootte van de Business Intelligence afdelingen waarbij opvallend is dat de grootte van de meeste Business Intelligence afdelingen 10 personen of minder is. Figuur 11 toont de functies zoals deze worden toegepast door Business Intelligence. Figuur 12 toont de aandachtsgebieden zoals deze worden toegepast binnen Business Intelligence.



Figuur 10 - Aantal personen werkzaam bij BI



Figuur 11- Toegepaste BI functies



Figuur 12 - Toegepaste aandachtsgebieden bij BI

Legenda figuur 11:

Rep = Reporting
 Ana = Analytics
 OLAP = Online analytical processing
 BPM = Business Performance Management
 DM = Datamining
 TM = Textmining
 CEP = Complex Event Processing

Legenda figuur 12:

MR = Management Rapportages
 CPM = Corporate Performance Management
 MB = Monitoring Bedrijfsactiviteiten
 OK = Optimaliseren Klantrelatie

4.1.4. Is Business Intelligence op de hoogte van het fenomeen Open Data?

Om te toetsen of Business Intelligence bekend is met het fenomeen Open Data is deze vraag voorgelegd in de survey. Bij een nee-antwoord is vervolgens gevraagd of men verwacht dat Open Data van toegevoegde waarde kan zijn.

Van de respondenten is 52 respondenten (77,6%) op de hoogte van het bestaan van Open Data. 15 respondenten (22,4%) waren nog niet op de hoogte van het bestaan. Van deze 15 respondenten denken 7 personen dat Open Data van toegevoegde waarde kan zijn voor de organisatie waar zij werkzaam zijn. 6 personen weten niet of het wel of niet toegevoegde waarde kan hebben. 2 personen denken dat het geen toegevoegde waarde heeft.

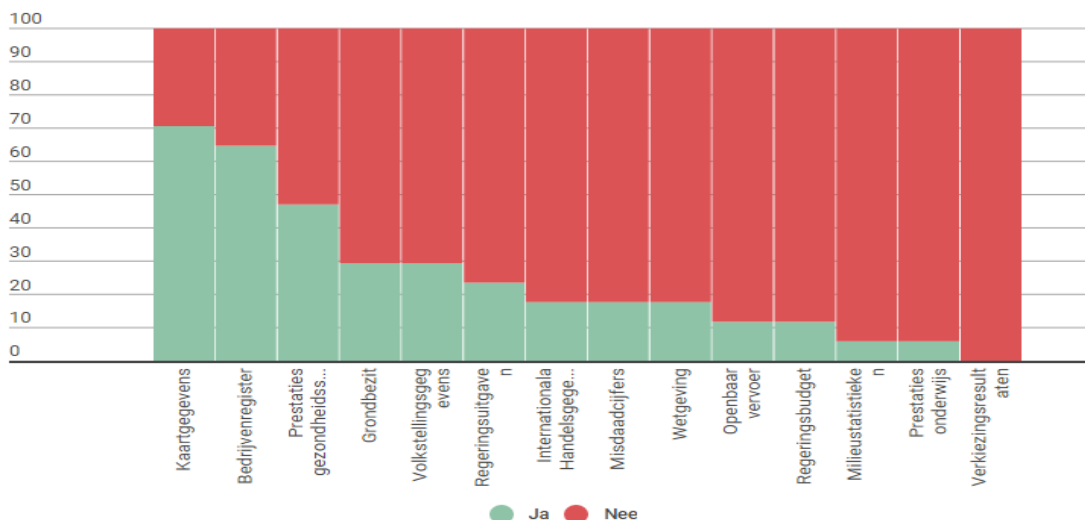
4.1.5. Maakt Business Intelligence gebruik van Open Data?

Van de 52 respondenten die bekend zijn met Open Data wordt door 17 organisaties (32,7%) daadwerkelijk gebruik gemaakt van Open Data. Door 35 organisaties (67,3%) wordt geen gebruik gemaakt van Open Data. Van de organisaties die gebruik maken van Open Data zijn 12 organisaties (70,6%) ervan overtuigd dat het gebruik een toegevoegde waarde heeft. 4 organisaties (23,53%) blijven neutraal over de toegevoegde waarde en 1 organisaties (5,88%) geeft aan dat Open Data geen toegevoegde waarde heeft.

Door business intelligence wordt bij het gebruik van Open Data vooral gebruik gemaakt van de vorm datasets. Alle respondenten geven aan dat van deze vorm gebruik wordt gemaakt. Voor wat betreft API, daar wordt in mindere mate gebruik van gemaakt. 47%, 8 van de 17 respondenten, geeft aan dat deze vorm van Open Data wordt gebruikt.

De Open Data bronnen die gebruikt worden, bestaan in alle gevallen uit overheidsdata. 70,6% van de respondenten geeft aan ook gebruik te maken van onderzoeksdata.

De categorieën Open Data die gebruikt worden door Business Intelligence worden weergegeven in figuur 13. Vooral data in de categorieën bedrijvenregister, kaartgegevens en prestaties gezondheidssector worden gebruikt.

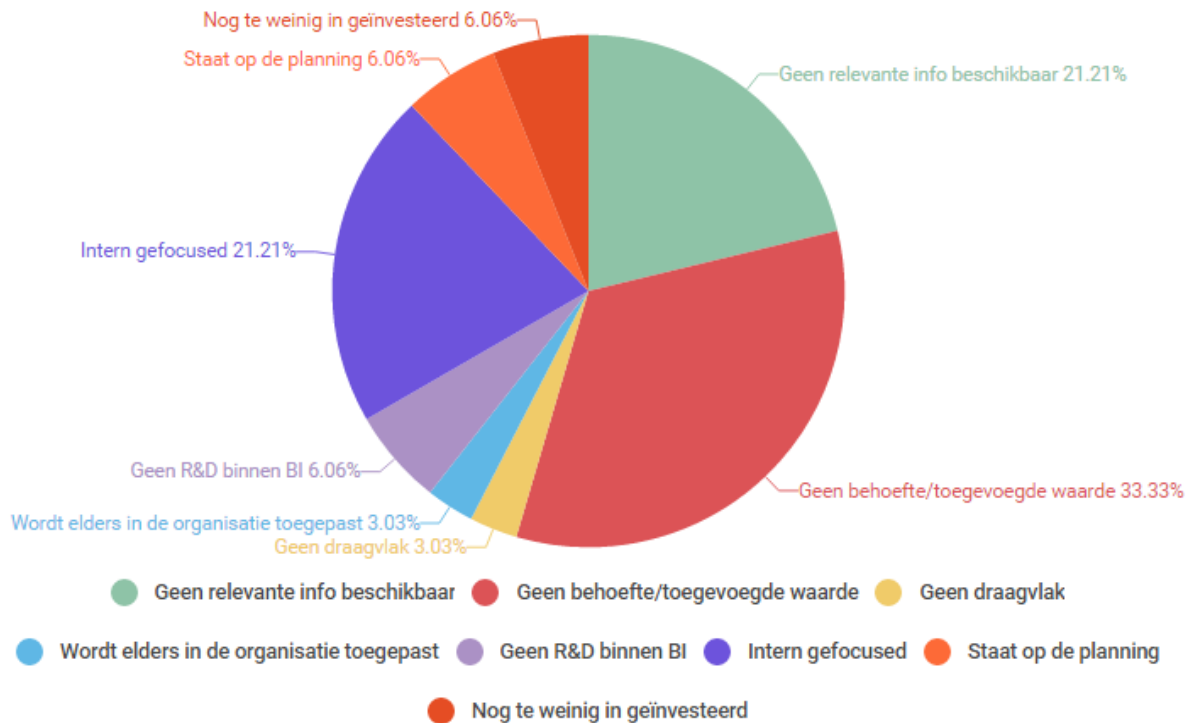


Figuur 13 - Overzicht van toegepast Open Data per categorie

Door de 35 respondenten die aangeven dat de Business Intelligence geen gebruik maakt van Open Data worden een aantal redenen gegeven waarom dat niet gebeurt:

- Er is niet direct Open Data beschikbaar die relevant is voor Business Intelligence.
- De organisatie waar de respondent werkzaam is, heeft geen behoefte aan toepassing van Open Data of ziet de toegevoegde waarde niet.
- Er is binnen de organisatie geen draagvlak om met Open Data aan de slag te gaan.
- Open Data wordt niet bij Business Intelligence maar elders in de organisatie toegepast.
- Er vindt geen *research & development* plaats binnen de BI organisatie.
- Business Intelligence is nog intern gefocust. Externe bronnen worden niet gehanteerd.
- De toepassing van Open Data staat op de planning.
- Er is nog te weinig tijd geïnvesteerd in de mogelijkheden van de toepassing van Open Data.

In figuur 14 volgt een grafiek van de reactie van de 35 respondenten.



Figuur 14 - Overzicht redenen van niet gebruiken Open Data

4.2. Meervoudige casestudy – Beantwoording E3 en E4

4.2.1. Cases

Er is enerzijds getracht cases voor de casestudy te vinden via het persoonlijke netwerk van de onderzoeker en anderzijds te werven via de uitgezette survey aan de top 500 private organisaties van Nederland. Bij de werving zijn 26 organisaties uit het netwerk van de onderzoeker benaderd. De organisatie **Achmea**, tevens werkgever van de onderzoeker, participeert in de casestudy. De andere organisaties hebben om uiteenlopende redenen niet geparticipeerd. Bij de werving via de uitgezette survey hebben twee organisaties aangegeven te willen participeren. Dat betreft een organisatie die **anoniem** wil blijven en de organisatie **Enexis**. De organisaties Achmea en Enexis, en de anonieme organisatie, hebben naast de casestudy geparticipeerd aan de survey.

4.2.2. Algemene kenmerken van de cases

De drie cases in de meervoudige casestudy worden beschreven in tabel 9. Bij de anonieme organisatie heeft één interview plaatsgevonden ten opzichte van zeven interviews bij Achmea en de zes interviews bij Enexis. De reden voor het enkele interview is dat de anonieme organisatie géén Open Data gebruikt. De volgorde van de cases is op chronologische volgorde van afname van interviews. De gehanteerde documenten uit de documentatiestudie van casestudy zijn opgesomd in bijlage I.

		Achmea	Enexis	Anoniem
Organisatie	Aantal Medewerkers (FTE)	17.000	4.300	1.900
	Omzet (in miljarden) 2014	20,2	1,4	0,8
	Sector	Financiële dienstverlener	Netbeheerder	Industrieel producent
	Werkzaam bij BI (FTE)	175	35	8
	Interviews	7	6	1

Tabel 9 - Eigenschappen cases

4.2.3. Case: Achmea

De organisatie Achmea is een financieel dienstverlener. Achmea is het moederbedrijf van een aantal Nederlandse verzekeringsmerken, zoals Centraal Beheer, Interpolis, FBTO, Zilveren Kruis en Avéro. Primair richt Achmea zich op de Nederlandse markt, maar slaat haar vleugels ook uit naar andere Europese landen. Binnen Achmea is de discipline Business Intelligence ondergebracht binnen de afdeling Competence Center Business Intelligence (CCBI). De afdeling heeft 175 FTE ondergebracht in tien subafdelingen. Er is een duidelijke scheiding zichtbaar naar afdelingen die zich bezig houden met beheer (Run) en afdelingen die zich bezig houden met ontwikkeling (Build). CCBI is een onderdeel van stafafdeling Informatie Management & Informatie Technologie (IM&IT).

4.2.3.1. Geïnterviewden

In de tabel 10 worden de respondenten getoond van Achmea. Vanwege de anonieme afname van de interviews worden de namen van de respondenten niet vermeld. In plaats daarvan wordt een unieke identificatie gebruikt. Bij het verwijzen naar respondenten is de hij-vorm gebruikt.

Identificatie	Functie	Afdeling	BI onderdeel?	Afname	Type interview
Achmea Respondent 1	Domeinarchitect	Strategy & Governance	Nee (IT)	Lync ⁴	Ongestructureerd
Achmea Respondent 2	PhD Student	Interpolis Marketing Particulieren	Nee (Business)	F2F ⁵	Ongestructureerd
Achmea Respondent 3	Data Scientist	Schade & Inkomen – Wheel of Knowledge	Nee (Business)	Telefoon	Ongestructureerd
Achmea Respondent 4	IT Specialist Analytics & Geo	Analytics & Geo	Ja	F2F	Semi gestructureerd
Achmea Respondent 5	Software Engineer ETL & GEO	Analytics & Geo	Ja	F2F	Semi gestructureerd
Achmea Respondent 6	Data Scientist	Analytics & Geo	Ja	F2F	Semi gestructureerd
Achmea Respondent 7	Data Scientist	Analytics & Geo	Ja	F2F	Semi gestructureerd

Tabel 10 - Respondenten Achmea

4.2.3.2. Beschrijving Business Intelligence

Aandachtsgebieden

Voor de aandachtsgebieden ‘management rapportages’ en ‘prestatie management’ worden ondersteund door Business Intelligence. Ook ‘optimaliseren klantrelatie en ‘standalone business intelligence applicaties’ worden toegepast, het zij in mindere mate. Dit wordt bevestigd door de respondenten 4 en 5.

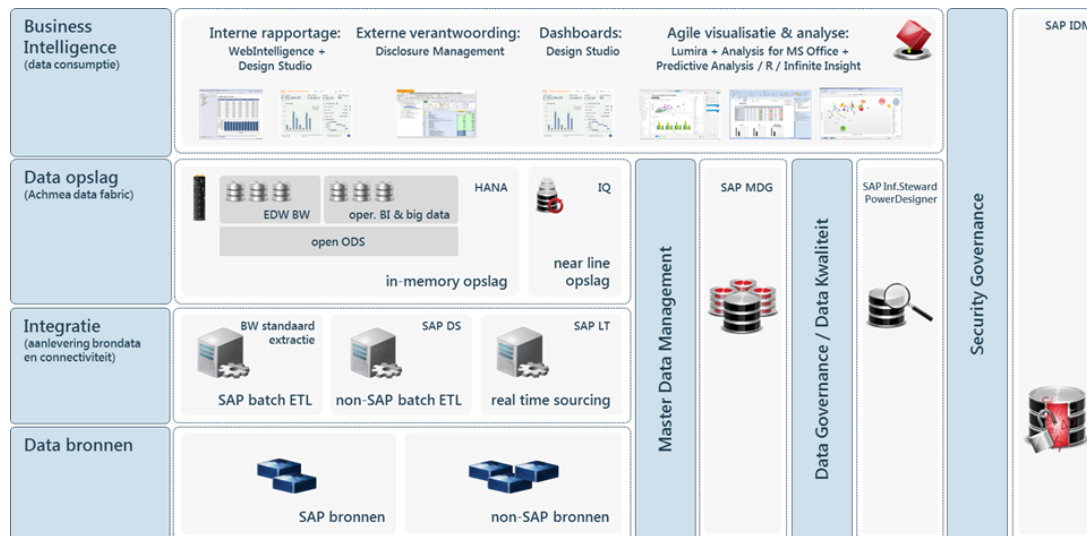
“Prestatiemanagement en management rapportages komen sterk naar voren. Gevolgd door optimaliseren klantrelatie, een voorbeeld is digitale klantbediening. Standalone BI applicaties komen ook voor, deze sluiten de lijst af.” (Respondent 5). Opvallend is ook dat ‘GEO’ als aandachtsgebied binnen BI een groot belang kent. *“Geo is een vreemde eend in de bijt bij Business Intelligence”* (Respondent 4).

⁴ Lync is een product van Microsoft. Het is een instantmessenger-, videogesprek- en telefonieprogramma. Het betreffende interview met respondent 1 is middels een videogesprek afgenomen.

⁵ F2F staat voor Face to Face

Architectuur

De Business Intelligence architectuur, zie figuur 15, binnen Achmea kan als traditioneel worden bestempeld. Databronnen, dataverplaatsing (integratie), data warehousing (Data opslag), middenlagen en front-ends komen sterk terug. Dit wordt bevestigd door respondent 5. Naast deze traditionele onderdelen van Business Intelligence komen Master Data Management, Data Governance/Data Kwaliteit en Security Governance terug. De BI architectuur binnen Achmea is SAP-dominant.



Figuur 15 - Architectuur Achmea

Functies

Alle bekende functies van Business Intelligence worden binnen Achmea toegepast. Sterk voorop staan Reporting en OLAP, gevolgd door Analytics. De toepassing van Analytics is sterk groeiende. Business Performance Management (BPM) wordt ook vaak toegepast. Datamining en Textmining worden ook toegepast, maar in mindere mate. Een niet-genoemde functie in de literatuur die wel aanwezig is binnen Achmea betreft de geografische analyse. De combinatie van Business Intelligence en Geografische Informatie Systemen wordt ook wel Location Intelligence genoemd. Location Intelligence focust op het verbeteren van processen en verschijnselen binnen een organisatie door bestudering van geografische relaties.

4.2.3.3. Beschrijving Open Data toepassing bij Business Intelligence

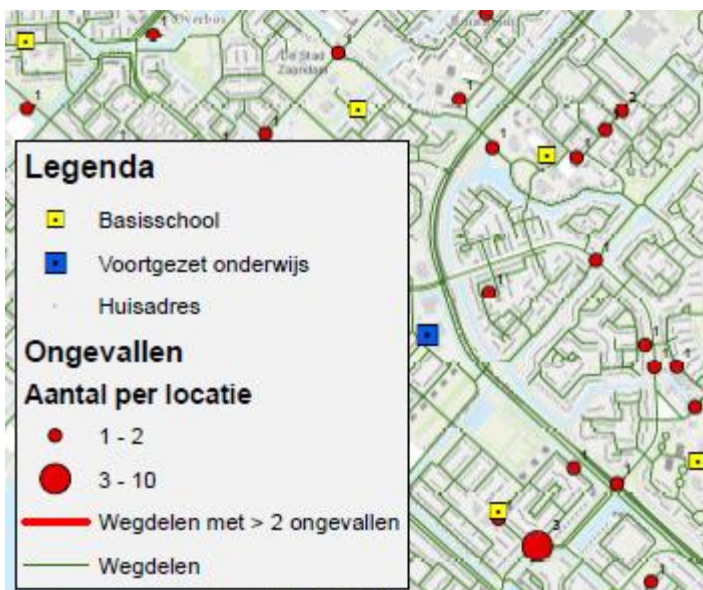
Verwantschap respondent met Open Data

Alle personen geïnterviewd bij de organisaties Achmea hebben tijdens hun dagelijkse werkzaamheden in min of meerdere mate te maken met Open Data. Respondent 1 is Domein Architect op het gebied van datamanagement. Op operationeel niveau maakt de respondent geen gebruik van Open Data. De respondent meldt dat de toepassing van verschillende databronnen, waaronder Open Data, vastligt in de Data Glossary. Ten tijde van het afnemen van het interview zijn 76 Open Data bronnen hierin opgenomen. Respondent 2, PhD student, is werkzaam binnen de business afdeling Marketing Particulieren. Hij maakt gebruik van Open Data. Daar waar nodig doet hij beroep op de kennis van personen binnen Business Intelligence. Respondent 3 voert Big Data projecten uit, waarbij regelmatig Open Data wordt gebruikt bij analyses. Respondent 4 wordt

gezien als de goeroe op Open Data gebied binnen Achmea. Hij past vooral Open Data toe bij het uitvoeren van geografische analyses. Respondent 5 is werkzaam op het gebied van ETL waarbij Open Data bronnen worden ingelezen in het datawarehouse. Respondent 6 voert met regelmaat onderzoeken en analyses uit. Daarbij maakt hij gebruik van Open Data. Respondent 7 is beperkt betrokken bij de Open Data toepassing binnen Achmea. Hij verneemt wat andere collega's met Open Data doen, en doet poging tot de ontwikkeling van Open Data tools in privétijd.

Doeleinden gebruik Open Data

Business Intelligence binnen Achmea maakt specifiek gebruik van Open Data om te kunnen visualiseren, communiceren, te verrijken en te kunnen voorspellen. De nadruk ligt daarbij op de toepassing van de geografische analyse. Een voorbeeld daarvan betreft een analyse naar het aantal ongevallen binnen een specifieke locatie, zie figuur 16.



Figuur 16 - Voorbeeld Open Data

Initiatief tot gebruik Open Data

Het initiatief tot gebruik van Open Data is buiten Business Intelligence ontstaan. Respondent 4 was werkzaam binnen de business van Achmea waar men ging werken met Geografische Informatie Systemen (GIS) en geodata. De belangstelling voor geo nam toe en door de steeds grotere verspreiding van Open Data door de overheid werd het gebruik bevorderd. Er ontstond een 'GEO-Team' in de business wat later werd opgenomen in het Competence Center Business Intelligence onder de vlag 'Analytics & Geo'. Daarmee is het gebruik van Open Data naar Business Intelligence verhuisd.

Invloed van Open Data op aandachtsgebieden Business Intelligence

Open Data heeft vooral invloed op het optimaliseren van de klantrelatie, geven de respondenten 4 t/m 7 aan. De aandachtsgebieden prestatie management, optimaliseren klantrelatie en standalone BI applicaties is veel meer op interne data gericht.

Respondent 5 geeft aan: *“Vooral op optimaliseren klantrelatie. Het wordt vaak gebruikt om snel op externe gebeurtenissen te reageren. Als er bijvoorbeeld ergens storm is geweest, kan men snel controleren welke risico’s er in de omgeving zijn.”*. Respondent 6 vult aan: *“Optimaliseren klantrelatie komt sterk terug bij Open Data. De klant neemt een product af, en hoe kunnen we dat product personaliseren. Vaak zie je dat er gegrepen wordt naar interne data en/of de combinatie met externe data.”*.

Invloed van Open Data op architectuur Business Intelligence

De invloed van Open Data op de Business Intelligence architectuur is vooral zichtbaar aan de kant van de databronnen. Er komen nieuwe externe databronnen bij. Respondent 5 geeft een voorbeeld: *“Bij twee applicaties worden de Open Data als externe bron ingeladen. De ene is Verras de Klant en de andere is Klant in Zicht. De gegevenssets gecombineerd en gevisualiseerd op kaarten.”*.

Naast de invloed op de externe databronnen houdt de Business Intelligence architectuur nu rekening met spatial data en –tooling. Dat zijn standalone Business Intelligence applicaties die beschikbaar worden gesteld.

Open Data heeft nog geringe invloed op andere onderdelen van de architectuur omdat men er vaak ad hoc mee aan de slag wil, en niet wil wachten totdat het geïmplementeerd is in een Business Intelligence omgeving. Respondent 6 geeft aan: *“Vaak wordt Open Data in de middenlaag geïntroduceerd. We gaan daar ontdekken wat we nodig hebben en op welke manier, waarna Open Data in een latere fase mogelijk in een datawarehouse worden opgeslagen. Vaak ga je ad hoc aan de slag met Open Data, en ga je niet lopen wachten totdat het een keer bij BI ingeladen is.”*.

Invloed van Open Data op functies Business Intelligence

De invloed van Open Data is vooral zichtbaar bij de functie Analytics. Daarbij zie je dat de nadruk vooral ligt op Predictive Analytics en in mindere mate bij Descriptive en Prescriptive Analytics. Bij Datamining en Textmining gaan wordt steeds vaker Open Data gebruikt.

Toepassing Open Data

Open Data wordt vooral ad hoc toegepast. Mensen die aan de slag gaan met Open Data, downloaden de data en slaan deze lokaal of op het netwerk op. Deze wordt dan verder gebruikt in Business Intelligence Self Service toepassingen of standalone BI applicaties. Bij structureel gebruik van datasets worden datasets automatisch verwerkt in het datawarehouse. Bij de toepassing van Open Data worden vaak de datasets van het Centraal Bureau van de Statistiek (CBS) gebruikt, de Basisregistraties Adressen en Gebouwen (BAG), het Koninklijk Nederlands Meteorologisch Instituut (KNMI) en het Actueel Hoogtebestand Nederland (AHN). Het CBS publiceert Open Data die inspeelt op de behoefte van de samenleving. BAG, beheert door het Kadaster, is de naam voor twee Nederlandse basisregistraties, de Basisregistratie Adressen (BRA) en de Basisgebouwenregistratie (BGR). De BAG gegevens herbergen de gegevens van alle gebouwen, objectklassen in Nederland, zoals woonplaats, nummeraanduiding en verblijfsobject. Het KNMI verzorgd gegevens over de weersverwachting. AHN is een digitale weergave van het aardoppervlak. Het is een bestand met gedetailleerde hoogtegegevens.

Een voorbeeld van een toepassing binnen Achmea die gebruik maakt van Open Data is de kaartapplicatie Verras de Klant (VdK). Met VdK kan bij een incident, bijvoorbeeld een aardbeving na gasboringen in Groningen, snel een ruimtelijke selectie worden gemaakt en de potentieel betrokken verzekerden adressen naar Excel worden geëxporteerd voor verder gebruik in het schadeproces. Door de applicatie wordt gebruik gemaakt van data uit BAG.

Een ander voorbeeld van de toepassing van Open Data betreft de ontwikkeling van een voorspellingsmodel waarmee de waarschijnlijkheid van inbraak te voorspellen is op huisnummerniveau. Dit voorspellingsmodel is opgezet met data van het CBS, KNMI en BAG.

Toegepaste vormen van Open Data bij Business Intelligence

Er worden vooral datasets gebruikt. Bij de uitvoering van een analyse werkt dat ook prettiger in vergelijking met API. Daarbij worden datasets vooral gebruikt bij ad hoc analyses, en vergt API een structurele inbedding. Het komt wel voor dat er op datasets nog bewerkingen moeten plaatsvinden voor gebruik. Respondent 5 geeft aan: *“Ik werk meestal met datasets waarop ik nog allerlei transformaties en verrijkingen uitvoer.”*. De respondenten geven aan dat

Gebruikte bronnen van Open Data bij Business Intelligence

Alle respondenten geven aan dat de voornaamste bron van Open Data de overheid is. Er worden regelmatig datasets opgehaald van ‘Open Data Nederland’, ‘Data.Overheid.NL’, PDOK (Publieke Dienstverlening op de Kaart) en Nationaal Geo Register. Er worden geen bronnen uit Open Research Data gebruikt. Hoewel de overgrote meerderheid van de Open Data afkomstig is van de Overheid, heeft respondent 4 ook aan: *“Vooral overheidsmateriaal wordt gebruikt. Daar heb je ook het stelsel van basisregistraties. Er wordt specifiek gebruik gemaakt van overheidsdata omdat er een kwaliteitscirkel aan gekoppeld zit. Fouten kunnen terug gemeld worden, worden verbeterd, etc. Betere data als overheidsdata kun je niet vinden. Het vertrouwd werken van Achmea houdt ook in dat er vertrouwde bronnen van data worden gebruikt.”*.

Categorieën Open Data in gebruik bij Business Intelligence

Er is Open Data beschikbaar in verschillende categorieën. De meeste Open Data die gebruikt wordt, betreft Open Data uit de categorieën kaartgegevens (Respondent 4, 5, 6 en 7), misdaadcijfers (Respondent 5, 6 en 7) en volkstellingsgegevens (Respondent 5, 6 en 7). Respondent 2 is specifiek bezig geweest met Open Data uit de categorie misdaadcijfers. Hij geeft aan: *“Het project is gestoeld op zowel Big als Open Data. Het doel van het project is de ontwikkeling van een voorspellingsmodel om de waarschijnlijkheid van inbraak te voorspellen op huisnummerniveau.”*.

Eisen ten aanzien van kenmerken van Open Data

In de tabel 11 is het belang ten aanzien van de kenmerken vastgelegd. De nadruk ligt bij Achmea op techniek. Het is vooral van groot belang dat Open Data toegankelijk, compleet en machine verwerkbaar is. Het minste belang wordt gehecht aan het non-discriminatoire zijn van Open Data.

Kenmerken Lindman	Kenmerken Borglund&Engvall en Ubaldi	R1	R4	R5	R6	R7	Score
Technisch	Toegankelijk	*****	*****	*****	****	****	23
	Compleet	*****	*****	*****	****	*****	24
	Machine verwerkbaar	****	*****	*****	***	****	21
	Tijdig	***	-	*****	***	****	15
	Primair	**	**	*****	***	**	14
Commercieel	Niet gebonden	****	**	**	**	*	11
	Non- discriminatoire	-	**	**	**	**	8
Wettelijk	Licentievrij	*****	*	**	**	****	14

Tabel 11 - Belangrijke kenmerken Achmea

Problemen bij de toepassing van Open Data

Het is van belang dat Open Data aan de kenmerken voldoet van toegankelijkheid, compleetheid en machine verwerkbaarheid. Kortgezegd moet de kwaliteit van een Open Dataset op orde zijn.

Naast de genoemde kenmerken is koppelbaarheid vaak een probleem. Het koppelen van Open Data met 'Closed Data' (Bedrijfsdata) is erg moeilijk. Open Data is vaak op plaats, wijk of postcode niveau beschikbaar, terwijl bedrijfsdata op huisnummer niveau beschikbaar is.

Een probleem dat vanuit de overheid wordt veroorzaakt is de grote hoeveelheid Open Data die wordt aangeboden. Respondent 6 geeft aan: *"Er is te veel Open Data aanwezig waardoor je door de bomen het bos niet meer ziet, anderzijds is de dataset die je zoekt vaak ook niet te vinden."*

4.2.4. Case: Enexis

De organisatie Enexis is een netbeheerder, voortgekomen uit Essent. Enexis distribueert elektriciteit en gas in Noord, Oost en Zuid Nederland. Enexis richt zich op de Nederlandse markt. Bij Enexis is de Business Intelligence discipline ondergebracht in afdeling Business Intelligence van de stafafdeling ICT. Binnen de afdeling zijn 35 FTE werkzaam, verdeeld over twee subafdelingen. Een afdeling Beheer en een afdeling Vernieuwing (Change).

4.2.4.1. Geïnterviewden

In de tabel 12 worden de respondenten getoond van Enexis. Net als bij Achmea zijn, vanwege de anonieme afname van de interviews, de namen van de respondenten niet vermeld. In plaats daarvan wordt een unieke identificatie gebruikt. Bij het verwijzen naar respondenten is de hij-vorm gebruikt.

	Functie	Afdeling	BI onderdeel?	Afname	Type interview
Enexis Respondent 1	BI Architect	BI Vernieuwing	Ja	F2F	Semi gestructureerd
Enexis Respondent 2	Datamanager / Projectleider	Asset Management	Nee (Business)	F2F	Ongestructureerd
Enexis Respondent 3	Database Marketeer	Klant & Markt	Nee (Business)	F2F	Ongestructureerd
Enexis Respondent 4	Informatie analist	BI Beheer	Ja	F2F	Semi gestructureerd
Enexis Respondent 5	Enterprise Architect	CIO Office	Nee (IT)	F2F	Ongestructureerd
Enexis Respondent 6	GIS System Analyst	GIS	Nee (Business)	F2F	Ongestructureerd

Tabel 12 - Respondenten Enexis

4.2.4.2. Beschrijving Business Intelligence

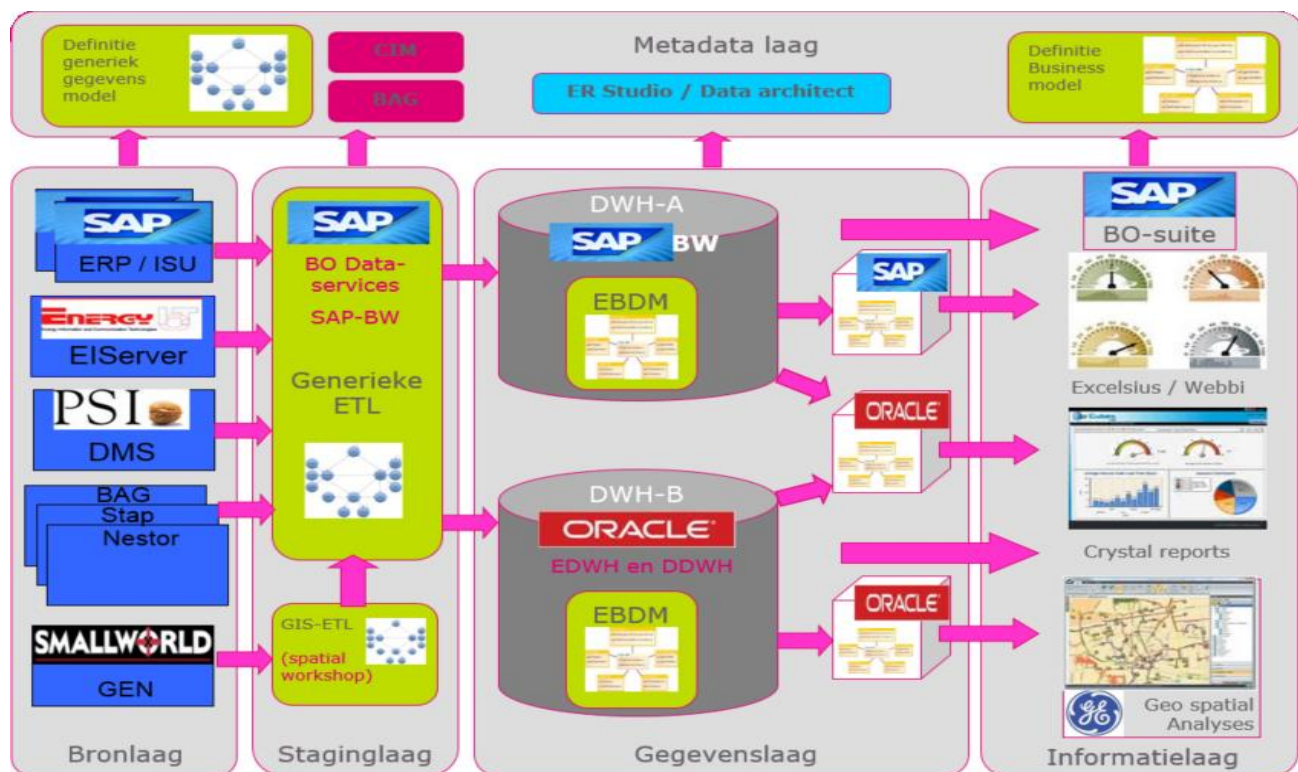
Aandachtsgebieden

De respondenten 1 en 4 zijn eensgezind over de aandachtsgebieden van Business Intelligence. De nadruk ligt op 'management rapportages' en 'prestatie management'. 'Optimaliseren klantrelatie' is in mindere mate van belang. "Optimaliseren klantrelatie doen we op zich ook, maar klanten zijn voor ons wat "minder van belang". (Respondent 1). Respondent 2 vult aan: "Optimaliseren klantrelatie gebeurt wel, al is dat kleinschalig. Dat heeft te maken met het bedrijf, we zijn geen commercieel bedrijf. We hebben klanten en proberen die klantrelatie ook te optimaliseren maar meer vanuit het besef dat Enexis een publieke functie heeft". Respondent 4 geeft aan dat naast de genoemde aandachtsgebieden het aandachtsgebied 'kwaliteitscontroles' veel door Business Intelligence wordt ondersteund. Hij geeft aan: "BI wordt vaak

gebruikt om datakwaliteit te bepalen en te meten. Dat gebeurt binnen Enexis veel. Je ziet dat redelijk wat data redelijk gedupliceerd is in het landschap. Daarom is er behoefte aan vergelijkende rapportages om te kijken of systemen consistent zijn aan elkaar”.

Architectuur

De Business Intelligence architectuur, zie figuur 17, is binnen Enexis klassiek te noemen. Naast het ontsluiten van verschillende databronnen, wordt ook data buiten de organisatie beschikbaar gesteld. Respondent 1 geeft aan: “Alle management informatie komt vanuit BI; rapportages, dashboards, data in ruwe vorm. We zorgen ook voor een stuk integratie waarbij we data uit verschillende bronnen, op basis van ETL, samenvoegen en beschikbaar stellen. Daar ga je verder als de traditionele BI afdeling. We stellen data extern beschikbaar, dus ook Open Data.”.



Figuur 17- Architectuur Enexis

Functies

Bij de organisatie Enexis worden alle functies ingezet. Wat net als bij Achmea een specifiek toegepast functie is, is de geografische analyse. Ook wel Location Intelligence genoemd, de combinatie tussen Business Intelligence en Geografische Informatie Systemen. Respondent 1 geeft aan: “Alle functies worden toegepast binnen Enexis. Wat echt uniek is binnen Enexis is de geografische analyse. Je zou ’t als Analytics willen zien, maar het past niet in de tooling en de omgevingen waardoor we het als aparte tak van sport zien.”.

4.2.4.3. Beschrijving Open Data toepassing bij Business Intelligence

Verwantschap respondent met Open Data

Alle personen geïnterviewd bij de organisatie Enexis hebben tijdens hun dagelijkse werkzaamheden in min of meerdere mate te maken met Open Data. Respondent 1 werkt op de subafdeling Vernieuwing van afdeling Business Intelligence. Als architect is hij bezig met innovatie, daarbij sterk bezig met Open Data met de centrale vraag: "hoe gaan we Open Data gebruiken en beschikbaar stellen aan de buitenwereld?". Respondent 2 is werkzaam als en Datamanager binnen de business van Enexis en vervult de rol van projectleider voor het project TEI (Toegankelijke Energie Informatie). Het project TEI is de basis voor de realisatie van een extern Open Data platform waar Open Data beschikbaar gesteld wordt. Het platform bestaat nog niet, men wil daarnaar toe gaan werken, het is een voornemen. Er worden verschillende cases uitgewerkt en er wordt een architectuuronderzoek uitgewerkt alvorens er wordt overgegaan op realisatie. Respondent 3 is database marketeer binnen de business van Enexis en maakt in zijn analyses gebruik van Open Data. Respondent 4 is Informatieanalist. Hoewel hij in de praktijk nog weinig doet met Open Data doet, gaat hij op korte termijn als Data Architect aan de slag waarbij hij ook zijn mening over Open Data moet gaan vormen. Respondent 5 is vanuit zijn rol als Enterprise Architect veel bezig met het thema data, de nadruk op Open Data. Hij is kartrekker achter het platform Open Data Netbeheer Nederland. Respondent 6 is vanuit zijn functie regelmatig bezig met de uitvoering van geografische analyses waarbij hij ook gebruik maakt van Open Data

Doeleinden gebruik Open Data

Naast het operationele doel, het verrijken van rapportages en het gebruik bij analyses, staat Open Data binnen Enexis als een van de strategische thema's opgesteld. De Open Data strategie is goedgekeurd door de directie met het gevolg dat binnen Enexis een 'Open Data Agenda' is opgesteld. In deze agenda staat beschreven wat Enexis de komende 3 jaren gaat doen op het vlak van Open Data. Dat uit zich in het uitvoeren van quick wins, uitvoering van projecten (Toegankelijke Energie Informatie, respondent 2) en het ontsluiten van Open Data.

Opvallend is ook dat Enexis niet alleen Open Data gebruikt, maar ook extern publiceert. Respondent 5 geeft aan: *"De data die wij publiceren, daarvan denken wij dat als partijen dit goed gaan gebruiken en wij dit actief volgen, er een steeds betere dialoog zal ontstaan. Wij hopen dat daardoor de marktpartijen die energiesysteem innovaties doen een beter inzicht krijgen in het totale energiesysteem en ons meenemen in hun inzichten. Zodanig kun je het totale systeem optimaliseren. Een andere reden is omdat wij een neutrale partij geacht worden te zijn. Geen commerciële partij, we willen geen stagnerende maar bevorderende factor zijn."* De openstelling van Open Data heeft een aantal redenen. Ten eerste wil Enexis de samenwerking met gemeenten en waterleidingmaatschappijen bevorderen waarbij respondent 1 het voorbeeld geeft: *"als er bijvoorbeeld ergens moet worden gegraven kunnen andere organisaties ook direct hun werk doen daar. "Als de patiënt toch open ligt"..."*. Ten tweede, Enexis ziet Open Data als een asset met de strategische doelstelling om de data publiekelijk te leveren. Respondent 1 geeft daarover aan: *"De data die we hebben, is erg gewild bij andere partijen. Het standpunt van Enexis is dat de data niet van ons is maar van de burger. Dat is toch nog een beetje het nutsbedrijf dat we zijn. Enexis is een private onderneming, er zit wel een zekere publiekelijke sfeer in de onderneming omdat de overheid aandeelhouder is."* De Open Data die Enexis momenteel aanbiedt, betreft de Kleinverbruikersdata.

Initiatief tot gebruik Open Data

Respondent 5 geeft het ontstaan van de toepassing aan van Open Data: *"Het initiatief binnen Enexis is ontstaan toen wij de SmartGrid roadmap ontwikkelde, de visie. Ik was een van de teamleden daarvan. Die is 2 jaar geleden afgerond, op verzoek van de directie. Smartgrid is een slim netwerk; met slimme meters en sensoren, allerlei ICT die je netwerk beter beheren en innoveren. In de smartgrid visie, leunend op 4 pijlers, stond in 1 pijler; Enexis gaat Open Data diensten realiseren. Toen was er nog geen idee hoe dat vormgegeven moest worden. Vanaf dag 1 werd dat geaccepteerd, er was niemand die zei; dat gaan we niet doen. Wat wel gevraagd werd; wat is het dan; leg het eens uit. Er zijn vervolgens projecten uitgekomen, waarbij we steeds meer geënthousiasmeerd werden door het onderwerp zelf. Toen kwamen we als snel met de gedachte; hier moet een business sponsor achter zitten. Toen zijn we een strategisch Open Data agenda gaan maken, wat inmiddels een strategische pijler is geworden."*

Invloed van Open Data op aandachtsgebieden Business Intelligence

Respondent 1 geeft aan: *"We kunnen preventief vervangingen van assets uitvoeren, dan scheelt dat ons én de burger miljoenen. Open Data levert dus niet alleen winst op, maar zorgt ook voor besparingen!"*. Het gebruik van Open Data heeft vooral effect op het aandachtsgebied 'kostenverlaging'. Een voorbeeld van kostenverlaging betreft het preventief vervangen van 'assets' op basis van Open Data. Assets zijn objecten die zich in het land bevinden waarvan Enexis de eigenaar is en waarvoor Enexis verantwoordelijk is. Voorbeelden van deze assets zijn middenspanningsstations, laagspanningsstations, gas- en elektriciteits-leidingen, sensoren, etc. Door toepassing van Open Data, het Actueel Hoogtebestand Nederland (AHN), kan een schatting worden gemaakt wanneer een asset moet worden vervangen, en wel op een moment voordat een storing kan optreden bij de asset. Met behulp van het AHN kan men bijvoorbeeld zien of een asset onder een busbaan of een gewone rijbaan ligt. Bevindt deze zich onder een busbaan, is de slijtage op de asset hoger en weet men dat er eerder vervanging moet plaatsvinden, waardoor eventuele schade wordt voorkomen. De preventieve vervangingen scheelt de organisatie miljoenen.

Invloed van Open Data op architectuur Business Intelligence

De invloed van Open Data op de architectuur is zichtbaar bij de frontends. De eindgebruikers van Business Intelligence producten maken gebruik van Self Service omgevingen waar ze zelf Open Data kunnen inladen voor de rapportage en/of de analyse.

Daarnaast heeft de externe publicatie van Open Data invloed op de architectuur. Daarvoor wordt bij Business Intelligence een datahub ingezet. Via de DMZ wordt Open Data beschikbaar gesteld aan externe partijen. Hier wordt van de klassieke Business Intelligence architectuur afgestapt. De Business Intelligence architectuur groeit naar een integratieomgeving en stapt meer af van de klassieke rapportageomgeving.

Aan de bronkant is lichte invloed merkbaar. Er worden nieuwe externe data geladen.

Invloed van Open Data op functies Business Intelligence

Men voorziet dat Open Data voornamelijk invloed heeft op Analytics en Datamining. Deze functies voorzien voornamelijk in het combineren van eigen data met externe data.

Toepassing Open Data

Bij Enexis wordt Open Data, net als bij Achmea, vooral ad hoc toegepast. Mensen die aan de slag gaan met Open Data, downloaden de data en slaan deze lokaal of op het netwerk op. Deze wordt dan verder gebruikt in Business Intelligence Self Service toepassingen of standalone BI applicaties. Bij structureel gebruik van datasets worden datasets automatisch verwerkt in het datawarehouse. Ook Enexis laat standaard de BAG-data in. Opvallend is dat Enexis, net als Achmea, gebruik maakt van dezelfde datasets. Bij de toepassing van Open Data worden de datasets van het Centraal Bureau van de Statistiek (CBS) gebruikt, de Basisregistraties Adressen en Gebouwen (BAG), het Koninklijk Nederlands Meteorologisch Instituut (KNMI) en het Actueel Hoogtebestand Nederland (AHN).

Een voorbeeld van de toepassing van Open Data is de usecase "Analyse Gasleiding". Het AHN is gebruikt om de hoogte van de grond boven leidingen te bepalen. Er wordt hierdoor een 3 dimensioneel beeld gecreeërd.

Toegepaste vormen van Open Data bij Business Intelligence

Er wordt zowel van datasets als van API gebruik gemaakt. De nadruk ligt zeer sterk op het gebruik van datasets. Datasets zijn snel bruikbaar, voor de gebruikmaking van API moeten structurele voorzieningen worden ingezet. Voor de publicatie van Open Data worden datasets beschikbaar gesteld.

Gebruikte bronnen van Open Data bij Business Intelligence

Open Data kan door iedere individueel en door iedere type organisatie, commercieel of niet-commercieel, (semi-)overheid of geen overheid, worden gemaakt en gepubliceerd. Uit de literatuur wordt duidelijk dat de nadruk bij publicatie van Open Data de bron vooral ligt bij onderzoek en bij overheden. Er wordt door alle respondenten aangegeven dat gebruik gemaakt wordt van Open Data uit overheidsbronnen. Respondent 1 geeft aan: "Vooral data van de overheid en semi-overheid wordt gebruikt."

Categorieën Open Data in gebruik bij Business Intelligence

Er is Open Data beschikbaar in verschillende categorieën. De meeste Open Data die gebruikt wordt, betreft Open Data uit de categorieën kaartgegevens (Respondent 1 t/m 6), milieustatistieken (Respondenten 2, 5 en 6), grondbezitgegevens (Respondenten 2, 3 en 6) en wetgeving (Respondent 2, 5 en 6).

Eisen ten aanzien van kenmerken van Open Data

In de tabel 13 is het belang ten aanzien van de kenmerken vastgelegd. De nadruk ligt bij Enexis op techniek. Het is vooral van groot belang dat Open Data actueel(tijdig), compleet, machine verwerkbaar en niet-gebonden is. Het minste belang wordt gehecht aan het licentievrij zijn van Open Data.

Kenmerken Lindman	Kenmerken Borglund&Engvall en Ubaldi	R1	R3	R4	R6	Score
Technisch	Toegankelijk	-	-	**	-	2
	Compleet	*****	-	*****	-	10
	Machine verwerkbaar	-	-	*****	****	9
	Tijdig	-	*****	****	****	13
	Primair	**	-	**	-	4
Commercieel	Niet gebonden	*****	-	-	****	9
	Non- discriminator	*	-	*	****	6
Wettelijk	Licentievrij	**	-	*	-	3

Tabel 13- Belangrijke kenmerken Enexis

Problemen bij de toepassing van Open Data

Bij de toepassing van Open Data worden verschillende problemen ervaren:

- **Bruikbaarheid.** Open Data is bij het eerste gebruik vaak niet verwerkbaar. Er moeten vooraf vaak nog kwaliteitsslagen overheen. Respondent 4 zegt hierover: *“De toepassing van Open Data zit nog vaak in de hobbyistensfeer en als je dan in allerlei bochten moet wringen om met die data te kunnen werken, is dat niet erg prettig.”*
- **Snelheid.** Vaak wil men snel over een dataset beschikken maar wordt dat beperkt door de trage levering.
- **Tijdigheid.** Men ervaart dat sommige datasets te laagfrequent worden ververs.
- **Koppelbaarheid.** Data is niet beschikbaar op huishoudniveau terwijl de eigen data op dat niveau aanwezig is. Daardoor moeilijk te koppelen.

4.2.5. Case: Anoniem

De anonieme organisatie betreft een industrieel producent. De organisatie richt zich vooral op de Nederlandse en Amerikaanse markt. De discipline Business Intelligence is binnen anonieme organisatie ondergebracht in de afdeling Business Intelligence van stafafdeling IT. Er zijn 8 FTE werkzaam op een enkele afdeling. Er zijn geen subafdelingen binnen de Business Intelligence organisatie. Daarbij is sprake van een combinatie beheer en change.

4.2.5.1. Geïnterviewden

In de tabel 14 wordt de respondent getoond van de anonieme organisatie. De naam van de respondent is niet vermeld. In plaats daarvan is een unieke identificatie gebruikt. Bij het verwijzen naar de respondent is de hij-vorm worden gebruikt.

	Functie	Afdeling	BI onderdeel?	Afname	Type interview
Anonieme Respondent 1	BI Manager	BI	Ja	Telefoon	Semi gestructureerd

Tabel 14 - Respondent anonieme organisatie

4.2.5.2. Beschrijving Business Intelligence

Aandachtsgebieden

De respondent van de anonieme organisatie geeft aan dat de focus van Business Intelligence volledig ligt op 'management rapportages'.

Architectuur

De Business Intelligence architectuur binnen de anonieme organisatie is als traditioneel te beschouwen. Het start met een Operational Data Store (ODS) van waaruit data in het data warehouse wordt geladen. Er worden vervolgens front-ends beschikbaar gesteld waar men management rapportages op kan draaien. Daarnaast wordt ook OLAP toegepast.

Functies

De anonieme organisatie maakt vooral gebruik van de functies Reporting en OLAP. In mindere mate wordt gebruikt gemaakt van Analytics en Business Performance Management. Datamining, Textmining en Complex Event Processing worden niet toegepast.

4.2.5.3. Beschrijving Open Data toepassing bij Business Intelligence

Verwantschap respondent met Open Data

De anonieme organisatie maakt geen gebruik van Open Data. Er is momenteel geen verwantschap van de respondent met Open Data. De organisatie maakt geen gebruik van Open Data omdat men vooralsnog geen toegevoegde waarde ziet in de toepassing. Voor de nabije toekomst zijn geen plannen om Open Data te gaan toepassen. De respondent is gevraagd over een aantal onderwerpen zijn verwachting uit te spreken.

Initiatief tot gebruik Open Data

Business Intelligence staat puur uitvoerend en ondersteunend opgesteld. Alle nieuwe activiteiten die gestart worden door Business Intelligence worden geïnitieerd vanuit de business.

Categorieën Open Data in gebruik bij Business Intelligence

De verwachting van de respondent is dat de volgende gegevenscategorieën van Open Data interessant zullen zijn voor gebruik; Kaartgegevens, Internationale handelsgegevens, Volkstellingsgegevens, Onderwijsprestaties, Bedrijvenregister en Wetgeving.

Eisen ten aanzien van kenmerken van Open Data

De respondent geeft aan dat toegankelijkheid en compleetheid waarschijnlijk de belangrijkste factoren zullen zijn. Non-discriminatoir en licentievrij worden als minst belangrijke factoren beschouwd.

Kenmerken Lindman	Kenmerken Borglund&Engvall en Ubaldi	Score
Technisch	Toegankelijk	5
	Compleet	5
	Machine verwerkbaar	3
	Tijdig	4
	Primair	3
Commercieel	Niet gebonden	3
	Non- discriminatoir	2
Wettelijk	Licentievrij	2

Tabel 15 - Belangrijke kenmerken anonieme organisatie

4.2.6. Overeenkomsten en verschillen bij de cases

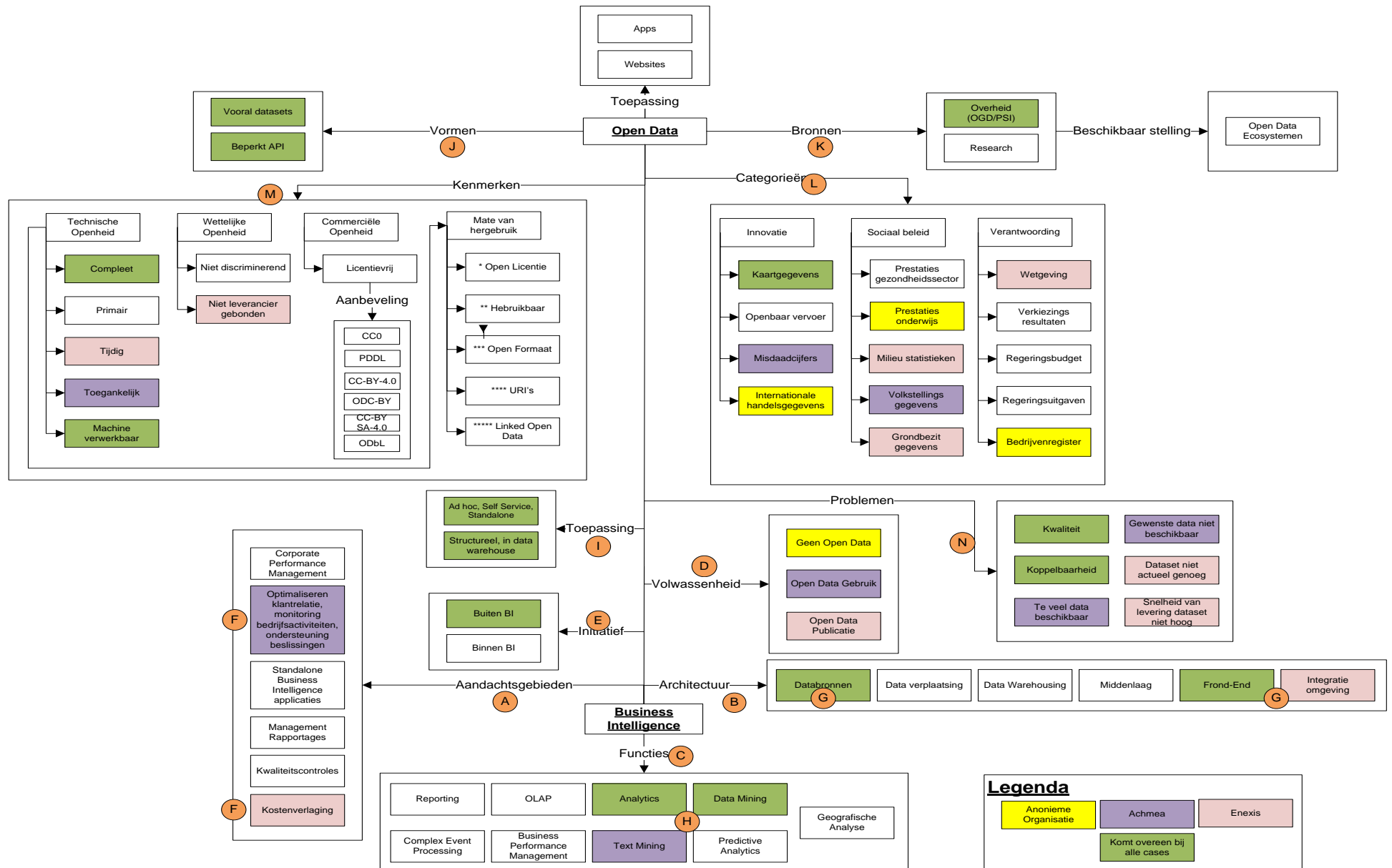
In tabel 16 wordt een overzicht gegeven van de overeenkomsten en verschillen per case, samenvattend 4.2.3. t/m 4.2.5.

		Achmea	Enexis	Anonieme organisatie
A	<i>Aandachtsgebieden BI</i>	- Managementrapportages - Prestatiemanagement	- Managementrapportages - Prestatiemanagement - Kwaliteitscontroles	- Managementrapportages
B	<i>BI Architectuur</i>	- Klassiek - SAP georiënteerd	- Klassiek - SAP georiënteerd	- Klassiek - Microsoft georiënteerd
C	<i>BI Functies</i>	- Worden allemaal gebruikt - Geografische analyse	- Worden allemaal gebruikt - Geografische analyse	- Alleen Reporting / OLAP
D	<i>Volwassenheid toepassing Open Data</i>	- Beperkt gebruik Open Data - Belegd binnen een afdeling	- Beperkt gebruik Open Data - Publiceert Open Data - Strategische pijler / Open Data agenda	- Geen toepassing Open Data
E	<i>Initiatief gebruik Open Data</i>	- Buiten BI	- Buiten BI	- Buiten BI
F	<i>Invloed Open Data op BI aandachtsgebieden</i>	- Optimaliseren klantrelatie	- Kosten verlagen	- N.v.t.
G	<i>Invloed Open Data op BI architectuur</i>	- Externe databronnen - Nieuwe standalone tooling	- Externe databronnen - Front-ends / Self-Service - Gaat fungeren als integratieomgeving	- N.v.t.
H	<i>Invloed Open Data op BI functies</i>	- Analytics - Datamining - Textmining	- Analytics - Datamining	- N.v.t.
I	<i>Toepassing Open Data</i>	- Ad hoc via Self Service en standalone Business Intelligence applicaties - Structureel in data warehouse	- Ad hoc via Self Service en standalone Business Intelligence applicaties - Structureel in data warehouse	- N.v.t.
J	<i>Toegepaste vormen van Open Data</i>	- Datasets - Beperkt API	- Datasets - Beperkt API	- N.v.t.
K	<i>Gebruikte bronnen van Open Data</i>	- Primair (semi-)overheid	- Primair (semi-)overheid	- N.v.t.
L	<i>Gebruikte categorieën van Open Data</i>	- Kaartgegevens - Misdaadcijfers - Volkstellingsgegevens	- Kaartgegevens - Milieustatistieken - Grondbezitgegevens - Wetgeving	- Kaartgegevens - Int. handelsgegevens - Volkstellingsgegevens - Onderwijsprestaties - Bedrijvenregister - Wetgeving
M	<i>Belangrijke kenmerken bij Open Data</i>	- Machine verwerkbaar - Compleet - Toegankelijk	- Machine verwerkbaar - Compleet - Tijdig - Niet gebonden	- Toegankelijk - Compleet
N	<i>Problemen bij toepassing Open Data</i>	- Kwaliteit Open Dataset - Koppelbaarheid - Te veel beschikbare data - Niet beschikbare data	- Kwaliteit Open Dataset - Koppelbaarheid - Snelheid - Tijdigheid	- N.v.t.

Tabel 16 - Overview overeenkomsten en verschillen tussen de cases

4.2.7. Overeenkomsten en verschillen per case geplot naar het referentiemodel

Onderstaande figuur 18 geeft een overzicht van de overeenkomsten en verschillen uit de resultaten van de casestudy geplot naar het referentiemodel uit de literatuurstudie. Recht onderaan in de figuur is de legenda opgenomen.



Figuur 18 - Overeenkomst/verschil geplot naar het referentiemodel

5. Conclusies en aanbevelingen

5.1. Conclusies literatuurstudie

De belangrijkste conclusie uit de literatuurstudie is dat er geen literatuur beschikbaar is over de toepassing van Open Data door Business Intelligence. Er is veel literatuur beschikbaar over de publicatie, het aanbod, van Open Data. De toepassing van, de vraag naar, Open Data wordt beperkt beschreven. Het is geen actief onderzoeksgebied.

L1: Wat is Open Data?

Uit de literatuurstudie blijkt dat er verschillende definities van Open Data zijn. Daarbij zijn er verschillende gerelateerde concepten te vinden als Open Government Data (OGD), Public Sector Information (PSI) en Big Data. Hoewel er een zekere overlap is tussen Open Data, OGD, PSI en Big Data zijn er toch wezenlijke verschillen. Zo zijn OGD en PSI data afkomstig uit de (semi-)overheid terwijl Open Data ook gepubliceerd kunnen worden door private ondernemingen of burgers. Big Data hoeft niet per sé open te zijn, en Open Data hoeft niet per sé Big te zijn.

Belangrijk is dat de Open kant van Open Data beschouwd kan worden als vrij toegankelijk, vrij te gebruiken, vrij te bewerken, vrij te delen voor ieder denkbaar doel. Uit de literatuurstudie is deze definitie voor Open Data afgeleid: *“Data die gebruikt kan worden, hergebruikt kan worden en opnieuw verspreid kan worden door iedereen – onderworpen enkel, in het uiterste geval, aan de voorwaarde dat men de bron vermeld en de data ook weer deelt”*.

Open Data is in verschillende vormen, datasets of API, beschikbaar. Daarbij zijn data in verschillende categorieën beschikbaar.

L2: Wie stelt Open Data beschikbaar?

Open Data is niet hetzelfde als OGD of PSI. Waarom raken zoveel mensen dan toch in verwarring over de term? Uit de literatuurstudie blijkt dat er één grote publicatiebron is. Dat betreft de publicatie van Open Data door (semi-)overheden.

Overheidsinstanties publiceren de Open Data om verschillende redenen. De nadruk ligt op het transparant maken van de werking van overheidsinstanties en de bevordering van de economische groei. Andere positieve bijeffecten zijn het efficiënter en effectiever maken van overheden, de bevordering van de toename van democratische controle en politieke deelneming en ter versterking van de wetshandhaving.

Om het gebruik tot Open Data te promoten zet de overheid een zogenaamd Open Data ecosysteem in. Dit ecosysteem moet er voor zorgen dat Open Data wordt gepubliceerd, vindbaar is, kwalitatief op een hoog niveau komt/blijft en dat er terugkoppeling kan plaatsvinden van eindgebruikers om diensten verder te verbeteren. Daarnaast worden aanvullende instrumenten ingezet om de toepassing van Open Data te bevorderen als educatie, conferences en zorgdragen voor open wetgeving.

L3: Welke kenmerken hebben Open Data?

Vanuit verschillende perspectieven worden de kenmerken van Open Data in de literatuur besproken. Zo wordt er gesproken over technische, wettelijke en commerciële openheid. Technische openheid behelst de kenmerken compleetheid, granulariteit, actualiteit, toegankelijkheid en machine verwerkbaarheid. Wettelijke openheid verwijst naar de licentie vrijheid en commerciële openheid gaat over het niet discriminerend zijn van die dienst en niet gebonden zijn aan *proprietary* leveranciers.

Tim Berners-Lee, vaak genoemd als grondlegger van het world wide web, geeft met zijn vijfsterrenmodel een weergave van het niveau van openheid bij Open Data. Het hoogste niveau van openheid in data, vijf sterren, wordt Linked Open Data (LOD) genoemd. LOD biedt interoperabiliteit van heterogene informatie bronnen.

L4: Hoe kan Open Data worden toegepast?

De nadruk in de beschikbare literatuur ligt op de aanbodzijde, de publicatie, van Open Data. De toepassing is voornamelijk zichtbaar in de opkomst van Websites & Apps die gebruik maken van Open Data. Daarbij is de toepassing vooral zichtbaar in de gemeenschapsdienstverlening.

Naast dit praktijkgebruik is in de literatuur beschreven welke verdienmodellen er toegepast kunnen worden vanuit het perspectief van het bedrijfsmodel.

De toepassing van Open Data is onderbelicht in de literatuur.

L5: Wat is Business Intelligence?

In de literatuur zijn veel verschillende definities vindbaar. Gekozen is om de definitie van Rud (2009) te hanteren: *“Een verzameling technieken en hulpmiddelen voor de omzetting van ruwe data naar betekenisvolle en bruikbare informatie voor business analyse mogelijkheden. Business Intelligence technologieën zijn in staat om grote hoeveelheden data te verwerken ter identificatie, ter ontwikkeling en ter ontwerp van nieuwe strategische kansen.”*.

De belangrijkste aandachtsgebieden van Business Intelligence zijn prestatie management, optimaliseren van de klantrelatie, standalone Business Intelligence applicaties en management rapportages.

L6: Hoe ziet een typische Business Intelligence architectuur er uit

De literatuur spreekt vaak over de ‘klassieke Business Intelligence architectuur’. Deze structuur bestaat uit databronnen, dataverplaatsing, datawarehousing, de middenlaag en front-ends.

L7: Welke functies kent Business Intelligence?

In de literatuur worden verschillende functies van Business Intelligence beschreven. Dat zijn Reporting, Online Analytical Processing (OLAP), Analytics, Datamining, Complex Event Processing, Business Performance Management en Textmining

5.2. Conclusies empirisch onderzoek

E1: Is Business Intelligence op de hoogte van het bestaan van Open Data?

Het empirisch onderzoek toont aan dat de bekendheid met het fenomeen Open Data bij Business Intelligence vrij hoog is. In de survey aan de top 500 private organisaties van Nederland is de vraag voorgelegd of men bekend is met het fenomeen Open Data. Daarbij is uiteengezet wat de definitie van Open Data is, en zijn enkele voorbeelden van Open Data sets gegeven. Het empirisch onderzoek toont aan dat de meeste Business Intelligence organisaties (77,6%) op de hoogte zijn van het bestaan van Open Data. Dat betreft 52 van de 67 respondenten. 22,4% van de respondenten, 15 personen, geven aan niet op de hoogte te zijn van het fenomeen.

E2: Maakt Business Intelligence gebruik van Open Data?

Het empirisch onderzoek toont aan dat door Business Intelligence beperkt gebruik gemaakt wordt van Open Data. In de survey is aan de respondenten (52) die bekend zijn met het fenomeen Open Data gevraagd of Business Intelligence gebruik maakt van Open Data. Door 35 organisaties (67,3%) wordt aangegeven dat geen gebruik van wordt gemaakt. Door 17 organisaties (32,7%) wordt dus wel gebruik gemaakt van Open Data.

Over het algemeen zijn organisaties die gebruik maken van Open Data tevreden over de toegevoegde waarde ervan. Vier organisaties blijven neutraal over de toegevoegde waarde en een organisatie geeft aan dat Open Data geen toegevoegde waarde heeft.

Wanneer men gebruik maakt van Open Data worden vooral datasets gebruikt en in mindere mate wordt gebruik gemaakt van API. Organisaties maken vaker ad hoc gebruik van Open Data dan dat er structureel gebruik van wordt gemaakt. Het gebruik maken van API wordt meer gezien als een structurele oplossing bij de toepassing van Open Data, er moet meer moeite gedaan worden om deze te gebruiken. Bij de toepassing van Open Data wordt vooral gebruik gemaakt van data uit overheidsbronnen.

Vooraf uit de categorieën bedrijvenregister, kaartgegevens en prestaties gezondheidssector wordt Open Data gebruikt.

Het niet toepassen van Open Data komt vooral voort uit het gebrek aan toegevoegde waarde, de beperkte externe focus van de Business Intelligence organisatie en het ontbreken van relevante data.

E3: Op welke manier past Business Intelligence Open Data toe?

De toepassing van Open Data door Business Intelligence kent verschillende volwassenheidsniveaus. Er kan sprake zijn van het niet toepassen van Open Data, het gebruiken van Open Data en het gebruiken én publiceren van Open Data. De anonieme organisatie past geen Open Data toe. Door Business Intelligence van Achmea wordt Open Data toegepast. Het staat nog in de kinderschoenen en concentreert zich binnen een van de subafdelingen, Analytics & Geo. De toepassing van Open Data bij Achmea vindt vooral op operationeel niveau plaats. Door Business Intelligence van Enexis wordt Open Data toegepast en gepubliceerd. De toepassing en de publicatie staat nog in de kinderschoenen. Open Data krijgt binnen Enexis veel aandacht door het te benaderen als een strategische pijler binnen de organisatie en de toepassing van Open Data

vanuit de 'Open Data Agenda'. In de literatuur zijn vele volwassenheidsmodellen bekend voor Business Intelligence, zoals het 'Business Information Maturity Model', 'TWDI's Business Intelligence Maturity Model', 'Gartner's Maturity Model for Business Intelligence and Performance Management', 'AMR Research's Business Intelligence/Performance Management Model, version 2', 'Business Intelligence Maturity Hierarchy' en 'Infrastructure Optimization Maturity Model' (Rajterič, 2010). Binnen de modellen zijn een verschillend aantal niveaus aanwezig met daarbij verschillende benamingen en typering per niveau. Overeenkomend in deze modellen is dat de toepassing van Business Intelligence met een niveau van 'onbekend' of 'initieel' start, en eindigt met een organisatiebrede 'continue optimalisatie' of 'doordrongenheid van belang'. De mate van volwassenheid in de toepassing van Open Data lijkt verbonden te zijn met de mate van volwassenheid van Business Intelligence.

Het initiatief tot het gebruik van Open Data ligt voornamelijk buiten de IT organisaties. De cases lieten zien dat het initiatief bij de business ligt.

De invloed van Open Data op Business Intelligence aandachtsgebieden is divers. De nadruk bij Achmea ligt voornamelijk op het optimaliseren van de klantrelatie. Open Data wordt ingezet om contact met de klant te verbeteren. Bij Enexis ligt de nadruk voornamelijk op kostenverlaging.

Afhankelijk van de Open Data strategie ontstaat impact op de Business Intelligence architectuur. Bij het gebruik maken van Open Data ontstaat vooral invloed op de externe databronnen. Bij de publicatie van Open Data zal de architectuur zich moeten gaan omvormen naar een integratieomgeving ter beschikbaarstelling van data. Het gebruik maken van Open Data heeft binnen Achmea invloed op de standalone tooling, terwijl er bij Enexis invloed ontstaat op de front-ends die als Self Service dienen waar Open Data ingeladen gaat worden.

Het gebruik van Open Data heeft vooral invloed op de Business Intelligence functies Analytics en Datamining. Dat wordt zowel door Achmea als Enexis ondersteund.

Facilitering tot het gebruik van Open Data door Business Intelligence vindt zowel plaats via self service en stand alone Business Intelligence applicaties als door het structureel opnemen van data in het data warehouse. Zowel Achmea als Enexis geven aan dat bij ad hoc gebruik van Open Data, dat wil zeggen wanneer een Open Data bron eenmalig of beperkt gebruikt wordt, gepoogd wordt dit op te lossen middels self service en stand alone Business Intelligence applicaties. Bij het structureel gebruiken van datasets wordt een permanente Business Intelligence oplossing gezocht voor de structurele opname in het data warehouse.

Bij het gebruik van Open Data wordt voornamelijk gebruik gemaakt van datasets en in beperkte mate wordt gebruik gemaakt van API. De cases Achmea en Enexis maken gebruik van datasets, minimaal gebruik van API.

Open Data van de overheid wordt primair als bron gebruikt. Bij het gebruik maken van Open Data maken Achmea en Enexis gebruik van data van de overheid.

Er wordt door Business Intelligence van verschillende categorieën Open Data gebruik gemaakt. De categorie kaartgegevens lijkt het meest gebruikt, verder verschilt het gebruik van categorieën per case. Waar Achmea vooral focust op kaartgegevens, misdaadcijfers en volkstellingsgegevens, focus Enexis vooral op kaartgegevens, milieustatistiek, grondbezitgegevens en wetgeving. De anonieme organisaties verwacht weer een andere set Open Data te zullen gebruiken.

Bij de toepassing van Open Data is een deels overeenkomende mening over het belang van de kenmerken van Open Data. Het machine verwerkbaar en compleet zijn lijken de belangrijkste kenmerken. Achmea geeft daarnaast aan dat het kenmerk toegankelijk erg belangrijk is, terwijl Enexis de nadruk legt op tijdig en niet-gebonden. De anonieme organisatie geeft aan, net als Achmea, dat de Open Data vooral toegankelijk moet zijn.

E4: Welke problemen worden ondervonden bij het gebruik van Open Data door Business Intelligence?

Bij de toepassing van Open Data door Business Intelligence worden voornamelijk externe problemen genoemd. Dat wil zeggen problemen die ontstaan bij de aanbieder van Open Data.

Het grootste probleem bij de toepassing van Open Data betreft het probleem dat de kwaliteit van de Open Dataset vaak te wensen over laat en dat het vaak niet inzichtelijk is wat de kwaliteit van een set is. Vaak moeten er verschillende kwaliteitsverbeteringen worden uitgevoerd om de Open Dataset werkbaar te maken. In de literatuur worden verschillende datakwaliteit dimensies genoemd: compleetheid, consistentie, uniciteit, validiteit en nauwkeurigheid (Askham et al., 2013). Deze dimensies van datakwaliteit zijn ook van toepassing op Open Data. Daarnaast ligt de nadruk bij literatuur over Open Data ook op technische, wettelijke en commerciële openheid.

Er is teveel Open Data beschikbaar. De overheid biedt momenteel dermate veel Open Data aan, dat men door de bomen het bos niet ziet. Het is niet gemakkelijk te overzien welke datasets allemaal beschikbaar zijn.

Tegenover de grote hoeveelheid Open Data staat het probleem van het niet beschikbaar zijn van Open Data. Vaak zijn specifieke datasets nodig voor een analyse waarbij de gevraagde dataset niet lijkt te bestaan.

Vaak is Open Data niet tijdig beschikbaar. Enerzijds wil men snel over Open Data beschikken, anderzijds wil men dat Open Data hoogfrequent beschikbaar wordt gesteld.

Tenslotte is er nog het probleem van moeilijke koppelbaarheid. In het kader van privacy wordt Open Data vaak niet tot op huisnummer niveau beschikbaar gesteld. Open Data wordt vaak op niveau van postcode, wijk, plaats of gemeente beschikbaar gesteld. Wanneer men eigen data wil combineren met Open Data ontstaat het probleem. Eigen data is vaak op huisnummerniveau beschikbaar. Enerzijds moeten hierdoor concessies worden gedaan (men wil eigenlijk op huisnummer niveau iets willen bepalen) en anderzijds bemoeilijk het technisch koppelen van de eigen data met de Open Data.

Bij de adoptie van nieuwe technologie, is het belangrijk dat een bepaald gebruiksgemak wordt ervaren. Het 'Technology Acceptance Model' is een theorie die modelleert hoe gebruikers een specifieke technologie accepteren en onderschrijft het belang van gebruiksgemak (Venkatesh & Bala, 2008). Bij de toepassing van Open Data is het belangrijk dat de gebruiker op een eenvoudige manier met de data kan werken.

5.3. Eindconclusie

De centrale vraag van het onderzoek is: ***In hoeverre is Business Intelligence binnen private organisaties zich bewust van het bestaan van Open Data, en op welke manier passen ze Open Data toe?***

Het aantal Business Intelligence organisaties dat zich bewust is van het bestaan van Open Data is hoog. Een beperkt aantal van die organisaties maakt daadwerkelijk gebruik van Open Data. De toepassing van Open Data, het initiatief komt vaak niet van Business Intelligence zelf, vindt ad hoc of structureel plaats. Daar waar ad hoc gebruik gemaakt wordt van Open Data worden door Business Intelligence stand alone applicaties en Self Service omgevingen geboden. Daar waar structureel gebruik gemaakt wordt van Open Data wordt de externe bron ontsloten, wordt Open Data geladen in het data warehouse en verder beschikbaar gesteld aan de middenlaag en de front-ends.

De toepassing van Open Data heeft op verschillende aandachtgebieden, architectuuronderdelen en functies van Business Intelligence invloed. Kenmerkend bij de toepassing van Open Data is dat vooral gebruik wordt gemaakt van datasets afkomstig van de overheid. Bij de toepassing worden verschillende categorieën Open Data aangeboord en houdt men rekening met het belang van de verschillende kenmerken.

Bij de toepassing van Open Data wordt met een aantal problemen rekening gehouden. De kwaliteit is vaak onder de maat. Er is te veel Open Data beschikbaar. Open Data mist, is niet beschikbaar. Open Data is niet tijdig beschikbaar. Open Data is niet gemakkelijk te koppelen met eigen data.

Bij de toepassing van Open Data is een mate van volwassenheid zichtbaar, variërend van het niet gebruiken van Open Data tot het gebruikmaken én publiceren van Open Data.

5.4. Aanbevelingen

Beschrijving van toepassing van Open Data

Over de aanbodzijde van Open Data is ontzettend veel geschreven in de literatuur. De vraagzijde, wie maakt er gebruik van Open Data, is zeer beperkt beschreven. Waarom wordt er in de literatuur zo weinig over de toepassing van Open Data geschreven?

Open Data bekendheid versus toepassing

Zoals te zien in de resultaten van de survey is 77,6% van de respondenten op de hoogte van het bestaan van Open Data. Daartegenover staat dat slechts bij 32,7% van de organisaties waar de respondenten werkzaam zijn, gebruik gemaakt wordt van Open Data door Business Intelligence. Daarbij zijn verschillende redenen gegeven waarom men er nog geen gebruik van maakt. Hoe kan in het algemeen bevordert worden dat Open Data verder gebruikt gaat worden? En hoe kan dit in het bijzonder voor Business Intelligence worden bevordert?

Open Data toepassing buiten Business Intelligence

Het onderzoek heeft specifiek plaatsgevonden naar de toepassing van Open Data door Business Intelligence. Over de toepassing van Open Data in zijn algemeenheid is momenteel nog weinig literatuur beschikbaar. De scope van het onderzoek kan verbreed worden: Op welke manier passen organisaties Open Data toe?

Toepassingsmogelijkheden Open Data naar branche

In de casestudy komt naar voren dat organisaties verschillende categorieën Open Data toepassen. Is er een verbijzondering mogelijk van Open Data categorie naar de branche waar een organisatie in werkzaam is? Daarmee kan mogelijk het probleem van de te grote hoeveelheid Open Data worden teruggebracht zodat Open Data makkelijker de juiste eindgebruiker kan gebruiken.

Volwassenheid van Open Data gebruik

Naar aanleiding van de uitvoering van de meervoudige casestudy is geconstateerd dat er een zekere volwassenheidsverschil is tussen de cases. Het ontwikkelen van een volwassenheidsmodel zou organisaties kunnen helpen om hun ambities met betrekking tot Open Data vorm te geven. Daarbij wordt aanbevolen te onderzoeken of er een verwantschap is tussen Business Intelligence volwassenheid en volwassenheid in de toepassing van Open Data.

Bevorderen Open Data publicatie

De (semi-)overheden zijn de grootste producenten van Open Data. Alle burgers en private/publieke organisaties kunnen Open Data publiceren. Wat zijn redenen voor organisaties om Open Data te publiceren? Welke problemen komen er kijken bij de publicatie? Een nieuw thema voor onderzoek.

5.5. Limitaties van het onderzoek

In deze paragraaf volgen een aantal punten van kritiek op het onderzoek.

Bij de uitvoering van de survey is een steekproefgrootte nodig van minimaal 217 respondenten om een juist externe validiteit te verkrijgen. Deze steekproefgrootte is bepaald op basis van een steekproefmarge van 5%, de populatie van 500 organisaties, een uitkomst van 50% en een betrouwbaarheidsniveau van 95%. Het verkrijgen van een respons van 43,4% (217/500) lijkt onmogelijk. Daarom is gestreefd naar een minimale respons van 50. Met de uiteindelijk respons van 67 wordt ruim voldaan aan de minimum respons. De externe validiteit is niet groot.

Bij het invullen van de survey is een navigatiepad gehanteerd, zie bijlage B. Respondenten die de vraag ODB1, "Is de Business Intelligence afdeling bekend met Open Data?", met nee hebben beantwoord, hebben de vervolgvraag, ODB2, gekregen: "De toepassing van Open Data zou voor de afdeling Business Intelligence waar ik werkzaam ben van toegevoegde waarde kunnen zijn?". Respondenten die op vraag ODB1 ja hebben geantwoord en vervolgens op vraag ODG1, "Maakt de Business Intelligence afdeling gebruik van Open Data?" ook ja, krijgen de vervolgvraag ODB3, "De toepassing van Open Data is van toegevoegde waarde voor de afdeling Business Intelligence?". Respondenten die wel van het bestaan van Open Data afweten, en waar de Business Intelligence afdeling geen gebruik maakt van Open Data is niet de vraag voorgelegd of zij de toepassing van Open Data als toegevoegde waarde beschouwen. De vraag over de toegevoegde waarde had aan alle respondenten voorgelegd moeten worden zodat een beeld verkregen had kunnen worden over het stuk van de populatie dat bekend is met Open Data maar het nog niet gebruikt.

Het onderzoek was exploratief van aard. In de literatuur is geen literatuur aangetroffen over de toepassing van Open Data door Business Intelligence. Bij de uitvoering van de literatuurstudie is er daarom voor gekozen om het separaat het gebied van Open Data en dat van Business Intelligence te ontginnen. Op basis van de literatuurstudie zijn twee samenvattende referentiemodellen ontwikkeld. Op basis van de referentiemodellen heeft de operationalisering plaatsgevonden en is de empirische studie uitgevoerd. In de literatuurstudie is geen documentatie gevonden over de publicatie van Open Data door private organisaties. Waar het onderzoek meer ging over het gebruik van de toepassing ontstond het nieuwe facet tijdens de tweede casestudy. De publicatie van Open Data door de private organisatie kwam dus als een verrassing.

6. Procesreflectie

Bij aanvang van het onderzoek is er onduidelijkheid geweest over de onderzoeksvragen en het onderzoeksthema. Er werd geen literatuur gevonden over de toepassing van Open Data bij Business Intelligence waardoor twijfel ontstond over de haalbaarheid van het onderzoek. Door de goede sturing van de begeleiders ben ik vol goede moed gestart met het onderzoek.

Het afstudeeronderzoek is uitgevoerd als onderzoeksproject met verschillende fasen. Aan het eind van iedere fase is een mijlpaaldocument opgesteld dat de fase heeft samengevat. Het proces is voorspoedig doorlopen en heeft naar planning ongeveer een jaar geduurd. In het begin van de afstudeerperiode is tweewekelijks gerapporteerd aan de begeleiders. Na verloop van tijd bleek dat toch te frequent en is dat afgezwakt. Daar waar heeft nodig heeft contact plaatsgevonden tussen onderzoeker en begeleider.

Het contact tussen de onderzoeker en de eerste begeleider was voornamelijk inhoudelijk van aard. Daar waar de onderzoeker vragen had, werden deze snel en adequaat beantwoord. Het werd de onderzoeker goed mogelijk gemaakt om zelfstandig het onderzoek uit te voeren. De tweede begeleider, tevens afstudeercoördinator, antwoordde snel op procesmatige vragen. Daarbij had zij belangrijke inbreng als meelezer. De begeleiding is als zeer positief ervaren.

De uitvoering van het empirisch onderzoek heeft meer tijd gevraagd dan gepland. De planning was 120 uur, naar schatting is uiteindelijk 200 uur besteed. Vooral het uitzetten van de survey en het transcriberen van interviews heeft meer tijd gekost dan gepland. Daarbij is alles achteraf getranscribeerd. Het is beter om na een interview de transcriptie te maken.

Een drukke baan, een gezin met twee kinderen en in de avonduren een onderzoek uitvoeren is een zware combinatie. Dat vergt een zekere mate van flexibiliteit en motivatie. Voor de uitvoering heeft de onderzoeker een vijftal vrije dagen beschikbaar gehad voor studie doeleinden. Daarnaast is het onderzoek, gepland 600 uren, verspreid over 50 weken. Er is een gemiddelde van 12 uren per week besteed.

De manier waarop wetenschappelijk onderzoek plaatsvindt, was nieuw voor de onderzoeker. Tijdens zijn HBO studie leerde hij hoe hij van punt A naar punt B moest komen en hoe hij moest ingrijpen op gebeurtenissen gedurende de tijd. Bij wetenschappelijk onderzoek ligt het begin en eindpunt in mindere mate vast en is methodiek zeer belangrijk. Dat vergt een volledig andere benadering.

De onderzoeker vond het leerzaam om het fenomeen Open Data te onderzoeken. Daarbij vond hij het prettig om een leemte in kennis op het toepassingsgebied te mogen vervullen.

Referenties

- Askham, N., Cook, D., Doyle, M., Fereday, H., Gibson, M., Landbeck, U., ... Schwarzenbach, J. (2013). *The Six Primary Dimensions for Data Quality Assessment*. Group, DAMA UK Working (p. 16).
- Bauer, F., & Kaltenböck, M. (2011). Linked open data: The essentials. *Edition Mono/monochrom*, Vienna.
- Bergsma, M. (2003). *Betrouwbaarheid en Validiteit van Kwalitatief geörienteerde Operational Audits* (pp. 1–39).
- BIAward. (2015). Wat is Business Intelligence? Retrieved from <https://www.biaward.nl/wat-is-business-intelligence-bi/>
- Bonina, C. (2013). New business models and the value of open data: definitions, challenges and opportunities. *Nemode.ac.uk*, (August issue).
- Borglund, E., & Engvall, T. (2014). Open data. Data, information, document or record? *Records Management Journal*, 24(2), 163–180.
- Bures, V., Otcenaskova, T., & Jasikova, V. (2012). The evaluation of external data resources for business intelligence applications: the example of the Czech Republic. *Journal of Systems Integration*, 3.1, 32–44.
- Cebotarean, E. (2011). Business intelligence. *Journal of Knowledge Management, Economics and Information Technology*, 1(2).
- Chakrabarti, S., Ester, M., Fayyad, U., & Gehrke, J. (2004). *Data mining curriculum: A proposal* (pp. 1–10).
- Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, 54(8), 88.
- Cirasella, J. (2014). Open Access. *Sage Publications*, 13(2), 12–14.
- Cohen, K., & Hunter, L. (2008). Getting started in text mining. *PLoS Computational Biology*, 4(1), e20.
- Coker, F. (2014). *PULSE: Understanding the vital signs of your business*. Ambient Light Publishing.
- Computerworld. (2006). *How Companies Are Implementing BI Competency Centers* (pp. 1–17).
- Davies, T. (2013). Open Data Barometer. *Open Data Institute and World Wide Web Foundation*, 1–45.
- Eckert, M., & Bry, F. (2009). Complex Event Processing (CEP). *Informatik-Spektrum*, 32(2), 163–167.
- Erkimbaev, a. O., Zitserman, V. Y., Kobzev, G. a., Serebrjakov, V. a., & Teymurazov, K. B. (2014). Publishing scientific data as linked open data. *Scientific and Technical Information Processing*, 40(4), 253–263.
- Fischer, S. (2014). Big Data: Herausforderungen und Potenziale für deutsche Softwareunternehmen. *Informatik-Spektrum*, 37(2), 112–119.
- Frolick, M. N., & Ariyachandra, T. R. (2006). Business Performance Management: One Truth. *Information Systems Management*, 23(1), 41–48.

- García-Peñalvo, F. J., García de Figuerola, C., & Merlo, J. a. (2010). Open knowledge: challenges and facts. *Web of Science*, 34(4), 520–539.
- Hester, J. R. (2014). Closing the data gap: Creating an open data environment. *Radiation Physics and Chemistry*, 95, 59–61.
- Hofstee, H., & Kusters, R. (2013). *Afstudeertraject Business Process Management and IT, introductie tot de cursus*.
- Hongyan, Y., & Mingdi, C. (2012). Reuse of Public Sectors Information in US: Challenges and Strategies to Information Resource Management. *Canadian Social Science*, 8(5), 83–90.
- Huijboom, N., & Broek, T. Van Den. (2011). Open data : an international comparison of strategies. *European Journal of ePractice*, 12(March/ April 2011), 1–13.
- Jansen, H. (2005). De kwalitatieve survey. *KWALON*, 30(3), 15–34.
- Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2014). Data-Driven Innovation through Open Government Data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 15–16.
- Jochems, M., & Joosten, R. (2005). *RU.nl - De gevalsstudie* (pp. 1–27).
- Kaasenbrood, M., & Zuiderwijk, A. (2015). Exploring the Factors Influencing the Adoption of Open Government Data by Private Organisations. *International Journal of Public Administration in the Digital Age*, 2(2), 75–92.
- Kahn, R., & Cannel, C. (1957). *The dynamics of Interviewing*. New York and Winchester.
- Kohavi, R., Rothleder, N. J., & Simoudis, E. (2002). Emerging trends in business analytics. *Communications of the ACM*, 45(8), 45–48.
- Kozak, M., & Hartley, J. (2013). Publication fees for open access journals: Different disciplines-different methods. *Journal of the American Society for Information Science and Technology*, 64(12), 2591–2594.
- Lindman, J. (2014). Similarities of Open Data and Open Source: Impacts on Business. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 59–70.
- Luhn, H. P. (1958). A business intelligence system. *IBM Journal of Research and Development*, 2(4), 314–319.
- Mackie, P., Sim, F., & Johnman, C. (2015). Big data! Big deal? *Public Health*, 1–2.
- Mayer-Schönberger, V., & Cukier, K. (2013). *De Big Data Revolutie* (6th ed.). Maven Publishing.
- Michalewicz, Z., Schmidt, M., Matthew, M., & Chiriack, C. (2007). *Evolutionary Computation in Dynamic and Uncertain Environments* (pp. 179–196). Springer Science & Business Media.
- Mijinyawa, M., & Abdulwahab, L. (2014). Evaluating the Adoption of Open Source Software. *British Journal of Applied Science & Technology*, 4(34), 4816–4834.
- Murray-Rust, P. (2008). Open Data in Science. *Serials Review*, 34(1), 52–64.

- Negash, S. (2004). Business intelligence. *The Communications of the Association for Information Systems*, 13(15), 176–195.
- Open Definition. (n.d.). Retrieved February 18, 2014, from <http://opendefinition.org/>
- Pflugfelder, E. (2013). Big data, big questions. *Communication Design Quarterly Review*, (August), 18–21.
- Rajterič, I. H. (2010). Overview of business intelligence maturity models. *Management: Journal of Contemporary Management ...*, 15(1), 47–67.
- Rud, O. (2009). *Business Intelligence Success Factors* (p. 284). John Wiley & Sons, Inc., Hoboken.
- Sandoval-Almazan, R. (2012). Open government 2.0: Citizen empowerment through open data, web and mobile apps. *Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance*, 30–33.
- Saunders, M., Lewis, P., Thornhill, A., Booij, A., & Verckens, J. (2011). *Methoden en technieken van onderzoek* (5th ed.). Pearson Education Benelux BV.
- Schijns, J., & Oude-de Wolf de, A. (2004). *Academische vaardigheden 3 - Observeren en rapporteren*. Open Universiteit, Heerlen.
- Standaard Bedrijfs Indeling 2008 Versie 2014*. (2014) (pp. 1–29).
- Tammisto, Y., & Lindman, J. (2012). Definition of open data services in software business. In *Lecture Notes in Business Information Processing* (Vol. 114 LNBIP, pp. 297–303).
- Tashakkori, A., & Teddlie, C. (2003). *Handbook of Mixed Methods in Social and Behavioural Research*. Thousand Oaks, CA: Sage publications.
- The Open Data Handbook. (2014). Retrieved December 17, 2014, from <http://www.opendatahandbook.org/>
- The Open Knowledge Foundation. (2014). Retrieved December 13, 2014, from <https://okfn.org/>
- Thoreson, K. (2013). Open Data Opens New Doors. *Public Management*, 95(October), 2013–2015.
- Ubaldi, B. (2013). Open Government Data - Towards Empirical Analysis of Open Government Data Initiatives. *OECD*, (22), 61.
- Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39(2), 273–315.
- Watson, H. J. (2014). Tutorial : Big Data Analytics : Concepts , Technologies , and Applications Tutorial : Big Data Analytics : Concepts , Technologies , and Applications. *Communications of the Association for Information Systems*, 34(1).
- Yin, R. K. (2009). *Case Study Research: design and methods* (4-th editi.). Sage publications.
- Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data : Essential elements of open data ecosystems. *Information Polity*, 19(1-2), 17–33.

Bijlagen

Bijlage A – Contacttemplates

Connectieverzoek LinkedIn

“Hallo <<voornaam>>, in mijn zoektocht naar respondenten voor mijn survey ihkv mijn afstuderen (Open Data en BI), vond ik jouw profiel op LinkedIn. Je zou me ontzettend kunnen helpen. BvD Joost van Poppel”

Verzoek tot invullen survey aan LinkedIn contact

Beste <<voornaam>>,

Bedankt dat je me als connectie hebt geaccepteerd! Heb je 5 minuten?

Naast mijn werk als Business Intelligence consultant bij Achmea, studeer ik in deeltijd aan de Open Universiteit, opleiding Business Process Management & IT. Momenteel ben ik aan het afstuderen en ben ik op zoek naar respondenten voor mijn survey en mijn casestudy.

De centrale vraag in mijn afstuderen is: “In hoeverre zijn Business Intelligence afdelingen binnen private organisaties zich bewust van het bestaan van Open Data, en op welke manier passen ze Open Data toe?”. Met Open Data wordt data bedoeld die vrij gebruikt kan worden, hergebruikt kan worden en opnieuw verspreid kan worden door iedereen (bijvoorbeeld gegevens van Kadaster, CBS, etc).

Mijn vraag jou:

Is het mogelijk dat jij, of 1 persoon van jouw BI-afdeling, een korte survey invult? Deze survey duurt 5 minuten en is anoniem. Idealiter is de persoon die de survey invult op de hoogte van de algemene gang van zaken van de BI-afdeling en (eventueel) van het fenomeen Open Data.

De survey kan in het Engels óf in het Nederlands worden uitgevoerd. De links:

Engels: <http://www.qbasiq.com/survey/index.php/389483/lang-en>

Nederlands: <http://www.qbasiq.com/survey/index.php/389483/lang-nl>

Ik zou je zeer erkentelijk zijn als je de survey kunt (laten) invullen. Als dank voor de participatie kan ik de uiteindelijke resultaten/thesis beschikbaar stellen.

Graag zie ik jouw reactie tegemoet.

Alvast bedankt!

Joost van Poppel

Verzoek tot invullen survey aan corporate website

Ter attentie van: Afdeling Business Intelligence

Onderwerp: Onderzoek naar gebruik Open Data door Business Intelligence afdelingen

Geachte heer, mevrouw,

Mijn naam is Joost van Poppel, werkzaam bij Achmea als BI Consultant. In deeltijd studeer ik aan de Open Universiteit, opleiding Business Process Management & IT, en ben bezig met het onderzoek voor mijn masterscriptie. In mijn onderzoek probeer ik antwoord te geven op de vraag: *“In hoeverre zijn **Business Intelligence afdelingen** binnen private organisaties zich bewust van het bestaan van **Open Data**? En op welke manier passen ze Open Data toe?”*. Middels een survey aan een aantal geselecteerde organisaties uit de top 500 organisaties van Nederland (Bron: Elsevier 2014), probeer ik een beeld te vormen over het bewustzijn van het bestaan van Open Data.

Uw organisatie, genoemd in de lijst van 500 grootste organisaties, wil ik verzoeken een korte survey in te vullen. Deze (anonieme) survey duurt **5 minuten**. Ik wil u verzoeken om één persoon op de afdeling aan te wijzen die de survey invult. Bij voorkeur is deze persoon bekend met de gang van zaken op de Business Intelligence afdeling en de gebruikte externe databronnen, voorbeelden: Manager, Enterprise Architect, Data Scientist, etc.

U kunt mij, en de wetenschap, ten zeerste helpen door te participeren in deze survey. Aan het eind van de survey wordt de mogelijkheid geboden om het emailadres achter te laten. De resultaten zullen dan digitaal aan u beschikbaar worden gesteld.

De survey kan gestart worden via de volgende link:

<http://www.qbasiq.com/survey/index.php/389483/lang-nl>

Hartelijk dank voor uw (re)actie!

Hoogachtend,

Joost van Poppel

joostvanpoppel@gmail.com

Bijlage B – Survey navigatiepad

Code	Groep	Vraag	Tonen indien
ALG1	Algemeen	Welke van de volgende functienamen beschrijft uw rol het beste?	-
ALG2	Algemeen	Hoe lang bestaat de Business Intelligence afdeling waar u werkzaam bent?	-
ALG3	Algemeen	Hoe langt bent u werkzaam bij de Business Intelligence afdeling?	-
ALG4	Algemeen	In welke branche is de onderneming waar u werkzaam bent actief?	-
BIA1	BI Afdeling	Hoeveel personen zijn werkzaam bij de Business Intelligence afdeling?	-
BIA2	BI Afdeling	Voor welke doeleinden wordt Business Intelligence binnen de onderneming ingezet?	-
BIA3	BI Afdeling	Welke Business Intelligence functies worden ingezet?	-
ODB1	Open Data Bestaan	Is de Business Intelligence afdeling bekend met het fenomeen Open Data?	-
ODB2	Open Data Bestaan	De toepassing van Open Data zou voor de afdeling Business Intelligence waar ik werkzaam ben van toegevoegde waarde kunnen zijn	ODB1 = Nee
ODG1	Open Data Gebruik	Maakt de Business Intelligence afdeling gebruik van Open Data?	ODB1 = Ja
ODB3	Open Data Bestaan	De toepassing van Open Data is van toegevoegde waarde voor de afdeling Business Intelligence?	ODG1 = Ja
ODG2	Open Data Gebruik	Van welke vormen van Open Data wordt op de Business Intelligence afdeling waar u werkzaam bent gebruik gemaakt?	ODG1 = Ja
ODG3	Open Data Gebruik	Van welke bronnen van Open Data wordt door de Business Intelligence afdeing gebruik gemaakt?	ODG1 = Ja
ODG4	Open Data Gebruik	Van welke categorieën van Open Data wordt door uw Business Intelligence afdeling gebruik gemaakt?	ODG1 = Ja
ODG5	Open Data Gebruik	Waarom maakt de Business Intelligence afdeling geen gebruik van Open Data?	ODG1 = Nee

Bijlage C – Coderingstabellen survey

Coderingstabel ALG1:

ID	Omschrijving	Type
ALG1_1	Architect	Radiobutton
ALG1_2	Bedrijfsanalist	Radiobutton
ALG1_3	Beheerder	Radiobutton
ALG1_4	Consultant	Radiobutton
ALG1_5	Informatiemanager	Radiobutton
ALG1_6	Manager	Radiobutton
ALG1_7	Projectleider / Scrummaster	Radiobutton
ALG1_8	Systeemspecialist	Radiobutton

Coderingstabel ALG2:

ID	Omschrijving	Type
ALG2_1	0 tot 1 jaar	Radiobutton
ALG2_2	Langer dan 1 jaar, korter dan 2 jaar	Radiobutton
ALG2_3	Langer dan 2 jaar, korter dan 3 jaar	Radiobutton
ALG2_4	Langer dan 3 jaar, korter dan 4 jaar	Radiobutton
ALG2_5	Langer dan 4 jaar, korter dan 5 jaar	Radiobutton
ALG2_6	Langer dan 5 jaar, korter dan 10 jaar	Radiobutton
ALG2_7	Langer dan 10 jaar	Radiobutton

Coderingstabel ALG3:

ID	Omschrijving	Type
ALG3_1	0 tot 1 jaar	Radiobutton
ALG3_2	Langer dan 1 jaar, korter dan 2 jaar	Radiobutton
ALG3_3	Langer dan 2 jaar, korter dan 3 jaar	Radiobutton
ALG3_4	Langer dan 3 jaar, korter dan 4 jaar	Radiobutton
ALG3_5	Langer dan 4 jaar, korter dan 5 jaar	Radiobutton
ALG3_6	Langer dan 5 jaar, korter dan 10 jaar	Radiobutton
ALG3_7	Langer dan 10 jaar	Radiobutton

Coderingstabel ALG4 (*Standaard Bedrijfs Indeling 2008 Versie 2014, 2014*):

ID	Omschrijving	Type
ALG4_1	Landbouw, bosbouw en visserij	Radiobutton
ALG4_2	Winning van delfstoffen	Radiobutton
ALG4_3	Industrie	Radiobutton
ALG4_4	Productie en distributie van en handel in elektriciteit, aardgas, stoom en gekoelde lucht	Radiobutton
ALG4_5	Winning en distributie van water; afval- en afvalwaterbeheer en sanering	Radiobutton
ALG4_6	Bouwnijverheid	Radiobutton
ALG4_7	Groot- en detailhandel	Radiobutton
ALG4_8	Vervoer en opslag	Radiobutton

ALG4_9	Logies-, maaltijd- en drankverstrekking	Radiobutton
ALG4_10	Informatie en communicatie	Radiobutton
ALG4_11	Financiële instellingen	Radiobutton
ALG4_12	Verhuur van en handel in onroerend goed	Radiobutton
ALG4_13	Advisering, onderzoek en overige specialistische zakelijke dienstverlening	Radiobutton
ALG4_14	Verhuur van roerende goederen en overige zakelijke dienstverlening	Radiobutton
ALG4_15	Openbaar bestuur, overheidsdiensten en verplichte sociale verzekeringen	Radiobutton
ALG4_16	Onderwijs of onderzoek	Radiobutton
ALG4_17	Gezondheids- en welzijnzorg	Radiobutton
ALG4_18	Cultuur, sport en recreatie	Radiobutton
ALG4_19	Overige dienstverlening	Radiobutton
ALG4_20	Huishoudens als werkgever	Radiobutton
ALG4_21	Extraterritoriale organisaties en lichamen	Radiobutton

Coderingstabel BIA1:

ID	Omschrijving	Type
BIA1_1	0 tot 10 personen	Radiobutton
BIA1_2	11 tot 25 personen	Radiobutton
BIA1_3	26 tot 50 personen	Radiobutton
BIA1_4	51 tot 100 personen	Radiobutton
BIA1_5	100 tot 200 personen	Radiobutton

BIA1_6	Meer dan 200 personen	Radiobutton
---------------	-----------------------	-------------

Coderingstabel BIA2:

ID	Omschrijving	Type
BIA2_1	Corporate Performance Management	Checkbox
BIA2_2	Optimaliseren Klantrelatie	Checkbox
BIA2_3	Monitoring Bedrijfsactiviteiten	Checkbox
BIA2_4	Management Rapportages	Checkbox

Coderingstabel BIA3

ID	Omschrijving	Type
BIA3_1	Reporting	Checkbox
BIA3_2	OLAP	Checkbox
BIA3_3	Analytics	Checkbox
BIA3_4	Data Mining	Checkbox
BIA3_5	Complex Event Processing	Checkbox
BIA3_6	Business Performance Management	Checkbox
BIA3_7	Text Mining	Checkbox

Coderingstabel ODB1, ODG1:

ID	Omschrijving	Type
ODB1_1	Ja	Checkbox
ODB1_2	Nee	Checkbox

Coderingstabel ODB2, ODB3:

ID	Omschrijving	Type
ODB2_1	Helemaal mee oneens	Checkbox
ODB2_2	Mee oneens	Checkbox
ODB2_3	Neutraal	Checkbox
ODB2_4	Mee eens	Checkbox
ODB2_5	Helemaal mee eens	Checkbox

Coderingstabel ODG2:

ID	Omschrijving	Type
ODG2_1	Datasets	Checkbox
ODG2_2	API	Checkbox

Coderingstabel ODG3:

ID	Omschrijving	Type
ODG3_1	Overheidsdata	Checkbox
ODG3_2	Onderzoeksdata	Checkbox

Coderingstabel ODG4:

ID	Omschrijving	Type
ODG4_1	Bedrijvenregister	Checkbox
ODG4_2	Gegevens grondbezit	Checkbox
ODG4_3	Gegevens openbaar vervoer	Checkbox

ODG4_4	Internationale handelsgegevens	Checkbox
ODG4_5	Kaartgegevens	Checkbox
ODG4_6	Milieustatistieken	Checkbox
ODG4_7	Misdadcijfers	Checkbox
ODG4_8	Prestaties gezondheidssector	Checkbox
ODG4_9	Prestaties onderwijs	Checkbox
ODG4_10	Regeringsbudget	Checkbox
ODG4_11	Regeringsuitgaven	Checkbox
ODG4_12	Verkiezingsresultaten	Checkbox
ODG4_13	Volkstellingsgegevens	Checkbox
ODG4_14	Wetgeving	Checkbox

Coderingstabel ODG5:

ID	Omschrijving	Type
ODG5_1	Redenen	Editveld

Bijlage D – Protocol casestudy

Hieronder volgt het te volgen protocol dat bij ieder interview zal plaatsvinden.

- a. Opening: De onderzoeker vertelt over de inhoud van het onderzoek (5 minuten)
- b. Kennismaking: De onderzoeker en respondent geven elkaars achtergrond aan (10 minuten)
 - a. Persoonlijke achtergrond
 - b. Werk achtergrond
 - c. School achtergrond
- c. Regels met betrekking tot de casestudy: (5 minuten)
 - a. U weet geen antwoord op de vraag: die wordt overgeslagen. Eventueel kan achteraf per email hierop gereageerd worden
 - b. Naar wens kan respondent anoniem blijven
 - c. Non-verbale reacties worden niet geregistreerd
 - d. Men kan zich zonder opgaaf van reden terugtrekken uit het onderzoek
 - e. Wanneer de respondent invloed / sturing van de onderzoeker ervaart tijdens het interview, dient dit direct gemeld te worden
 - f. Achteraf wordt de respondent gevraagd de samenvatting van het interview te accorderen.
- d. Interviewvragen Semigestructureerd, niet te gebruiken bij ongestructureerd (50 minuten):
 1. *Kunt u de afdeling omschrijven waar u werkzaam bent?*
 - a. *Positionering binnen de organisatie*
 - b. *Aantal medewerkers*
 - c. *Organigram*
 - d. *Aandachtsgebieden van BI*
 - e. *Architectuur van BI*
 - f. *Functies van BI*
 2. *Wat is volgens u de definitie van Open Data?* - Vervolgens de definitie uit de literatuur bespreken en op een lijn komen met de respondent.
 3. *Wat is volgens u de definitie van Business Intelligence?* – Vervolgens de definitie uit de literatuur bespreken en op een lijn komen met de respondent.
 4. *Welke vormen van Open Data worden binnen uw afdeling gebruikt?* – Vervolgens de vormen uit de literatuur bespreken.
 5. *Welke categorieën Open Data worden binnen uw afdeling gebruikt?* – Vervolgens categorieën uit de literatuur vermelden.
 - a. *Met welk doel worden deze categorieën gebruikt?*
 - b. *Waarom worden andere categorieën niet gebruikt?*
 6. *Welke bronnen van Open Data worden binnen uw afdeling gebruikt?* – Vervolgens bronnen uit de literatuur vermelden.
 - a. *Worden andere bronnen gebruikt? Zo ja, welke?*
 7. *Hoe wordt Open Data door de binnen uw afdeling ontsloten vanuit de leverende partij?*
 - a. *Wordt Open Data handmatig gedownload? Geautomatiseerd?*
 - b. *Zijn er specifieke handmatige / geautomatiseerde processen ingericht?*
 8. *Hoe gaat uw afdeling om met de verschillende kenmerken van Open Data?* – Vervolgens kenmerken uit de literatuur vermelden.
 - a. *Per kenmerk bespreken*

- i. *Is dit kenmerk bevorderend of beperkt ten opzichte van het gebruik van Open Data, en waarom?*
 - ii. *Hoe gaat de afdeling om met de kenmerken van het gebruik van Open Data, en waarom? (Bijvoorbeeld: tijdigheid van gegevens)*
 - 9. *Met welke doeleinden maakt de afdeling gebruik van Open Data? Wie neemt initiatief?*
 - 10. *Hoe is het initiatief tot het gebruik van Open Data ontstaan?*
 - 11. *Binnen welke BI aandachtsgebieden worden Open Data toegepast – Vervolgens aandachtsgebieden uit de literatuur vermelden.*
 - 12. *Op welke onderdelen uit de BI architectuur hebben Open Data impact? – Vervolgens BI architectuur uit de literatuur bespreken.*
 - 13. *Op welke BI functies hebben Open Data impact? – Vervolgens BI functies uit de literatuur bespreken.*
- e. Verzoek tot documentatie: de onderzoeker vraagt de respondent of hij documentatie beschikbaar heeft over de toepassing van Open Data. Daarbij zal aangegeven worden dat na de afronding van het afstudeeronderzoek de documenten uit de bestandsopslag van de onderzoeker verwijderd zullen worden.
- f. Afsluiting interview:
- a. De onderzoeker vraagt de respondent aan te geven wat goed ging in het interview, en wat beter kon. Dit zal in opvolgende interviews worden meegenomen.
 - b. De onderzoeker geeft nogmaals aan dat deze de samenvatting zal toesturen. Eventueel emailadres kan hier gevraagd worden.
 - c. De onderzoeker bedankt de respondent tot deelname
- g. Na het interview en de levering van documentatie zullen per document de volgende acties plaatsvinden:
- a. Het interview wordt gedocumenteerd
 - b. Het document wordt opgeslagen in de gegevensopslag
 - c. De geleverde documentatie wordt doorgenomen en samengevat. Daarbij worden de volgende eigenschappen beschreven:
 - i. Doel van ieder document
 - ii. De omstandigheden van de totstandkoming van het document
 - iii. Doelgroep van documentatie
 - iv. Aanwezigheid bevestigend of ontkrachtende documentatie
 - d. De gegevens in de documentatie worden georganiseerd en gecombineerd tot visuele netwerkweergaven. Deze visuele weergave geeft een beeld van de inhoud van de documentatie door gebruik te maken van een verzameling knooppunten en/of kaders die verbonden of gekoppeld zijn door lijnen om verbanden aan te geven.
 - e. Het interview wordt uitgeschreven en de respondent ontvangt in een document beschrijving van het interview, samen met de samenvatting van de doorgenomen documentatie.

Bijlage E – Correspondentie auteurs Open Data artikelen

De aanschrijving als volgt:

Dear <<naam auteur>>,

I recently read your article (<<Naam artikel>>) and found it very interesting to read. I read the article in the context of my thesis research. The principal research question I want to answer is: "To what extent is Business Intelligence within private companies aware of the existence of Open (Government) Data, and if they are aware, how do they utilize them?".

In my literature study I have not been able to find a single article that discusses the utilization of Open Data in a Business Intelligence environment.

As I want to make sure that there are no scientific articles available on the subject, I'm contacting you. My question for you is: Do you know of any research regarding the use of Open Data in Business Intelligence environments?

Auteur		Reactie
Erik Borglund	Associate professor, PhD, Mid Sweden University, Faculty of Science, Technology and Media, Department of Archives and Computer Science	I am happy that someone read my article:) But when it comes to your question I am unfortunately unaware about any research in the business intelligence field. However I believe that "BI" will have a chance to use open data. I believe they are aware but that the uniqueness of the data is nothing you find comments about in research
Thorhildur Jetzek	Open data analyst/PhD student at KMD, Denemarken	<p>You are right, there is not much research on use of Open Government Data in private sector companies, but that makes it an exciting new area of research :) So far, most of the research I have found remotely related to that area is either on a) Linked data methods, mostly published in Computer Science outlets OR b) Business models for data intermediaries</p> <p>Here is one paper that discusses use of OGD in BI context in a public sector organization. I could not retrieve the text from our CBS library so I send you the link. http://link.springer.com/chapter/10.1007/978-3-319-10178-1_7</p> <p>The attached paper might be more helpful as it also summarizes the literature and looks at Dutch companies. I got it from one of the authors, she did not have the final published version so this is a "draft version" :)</p> <p>Best of luck with your thesis, is it a PhD or a M.Sc.? I would like to see it when you are finished, if that is OK :)</p>

Michael Kaschesky	Project Director (PhD/MBA) Swissdat	<p>Good question! Aware yes, but clueless on how to utilize open gov data. One big stumbling block is the heterogenous quality of open data not just in terms of formatting but also and more importantly in terms of relevance. Aggregated yearly deaths per city district may not be interesting for everyone. What is being published differs widely. But there are interesting use cases, such as using open data about restaurants, museums, and events to recommend a neighbourhood for living.</p> <p>A few weeks ago during Spaghetti Open Data meeting 2015 I came across Letizia Pernigotti and Alessandro Portosa from SpagoBI which uses open data in their BI tools. Perhaps check with them would be a good idea [www.spagobi.org].</p> <p>There's a workshop on Visualisation in Business Intelligence and Open Data in Barcelona 21-24 July [http://www.graphicslink.co.uk/IV2015/VBI_OD.htm]</p> <p>So definitely an upcoming and interesting topic that requires more time and work on both sides, gov and biz.</p>
Stefan Kulk	Promovendus, department Rechtsgeleerdheid, expertise auteursrecht / intellectueel eigendomsrecht.	Ik ken geen literatuur m.b.t. het op de hoogte zijn van open data
Juho Lindman	Assistant Professor of Information Systems, Department of Management and Organization, Hankel School of Economics. Helsinki, Finland.	I really do not know any work that would look open data specifically from BI perspective. However, I would speculate that this is because the openness/closedness of the actual data source is not as big issue in BI as is the reliability and long-term availability of the said data. The API-viewpoint however would I think be interesting to have a look also from BI perspective. Keep me posted on your research and good luck in pursuing the topic.
Peter Parycek	Head of Center E-Governance. Danube-University Krems	Thanks for your feedback. You are raising an important question, which is not answered yet based on my knowledge. It's highly relevant and I mention that topic in my talks, so it would be great if I could reference on your thesis.
Johann Höchtl	Research fellow at the center of E-Governance	Organisational side: I think for most decision makers, the "Veracity" in Big Data is still the limiting factor for widespread adoption of open data in BI solutions. Basing

		<p>your decisions on data of questionable quality seems like a dangerous play.</p> <p>This scholarly article might be of your interest: http://dl.acm.org/citation.cfm?id=2320812</p> <p>Technical side: Another issue in using open data as a basis for BI is that the data is not available using APIs on a cloud service but you are rather responsible on your own to download the data, make it available using eg. HDFS, build your own DW and on top of that your OLAP BI cubes. Would open data made available on cloud infrastructure, exposed through APIs which would allow virtualized integration on a DFS of what kind ever, the situation would be easier.</p>
Marijn Janssen	Professor in "ICT and Governance", Delft University	Often the term data analytics is used instead of business intelligence. Look at this term in relationship to open data. Also in our Engage project we used all kind of business intelligence mechanisms, although we did not use these terms. I hope this is of help.
Tim Davies	Independent researcher: open data.	<p>Nothing immediately springs to mind, though this might be a good question to ask at: https://www.linkedin.com/groups?home=&gid=4790214. There have been some studies by Deloitte and McKinsey on the economic value of open data which might tangentially touch upon this. E.g. in http://www2.deloitte.com/uk/en/pages/deloitte-analytics/articles/open-data-driving-growth-ingenuity-and-innovation.html</p>
Luigi Selmi	Research Associate, Bern University of Applied Sciences	big companies are aware of Open (Government) Data and have been using data sets provided by public administration or organizations worldwide for their businesses. Google uses transport data published by public administration for their Google Transit service, many companies have been using weather forecast data, Insurance companies use statistics data sets (Eurostat, Census.gov) for risk assessment. That said I can't remember a research paper about the use of Open Data in Business Intelligence environments in general. It would be very interesting to know what data sets they use, how they process them and what kind of services they provide based on those data sets. Maybe you already know ODI

and OD500 that are conducting research about the use of Open Data for businesses.

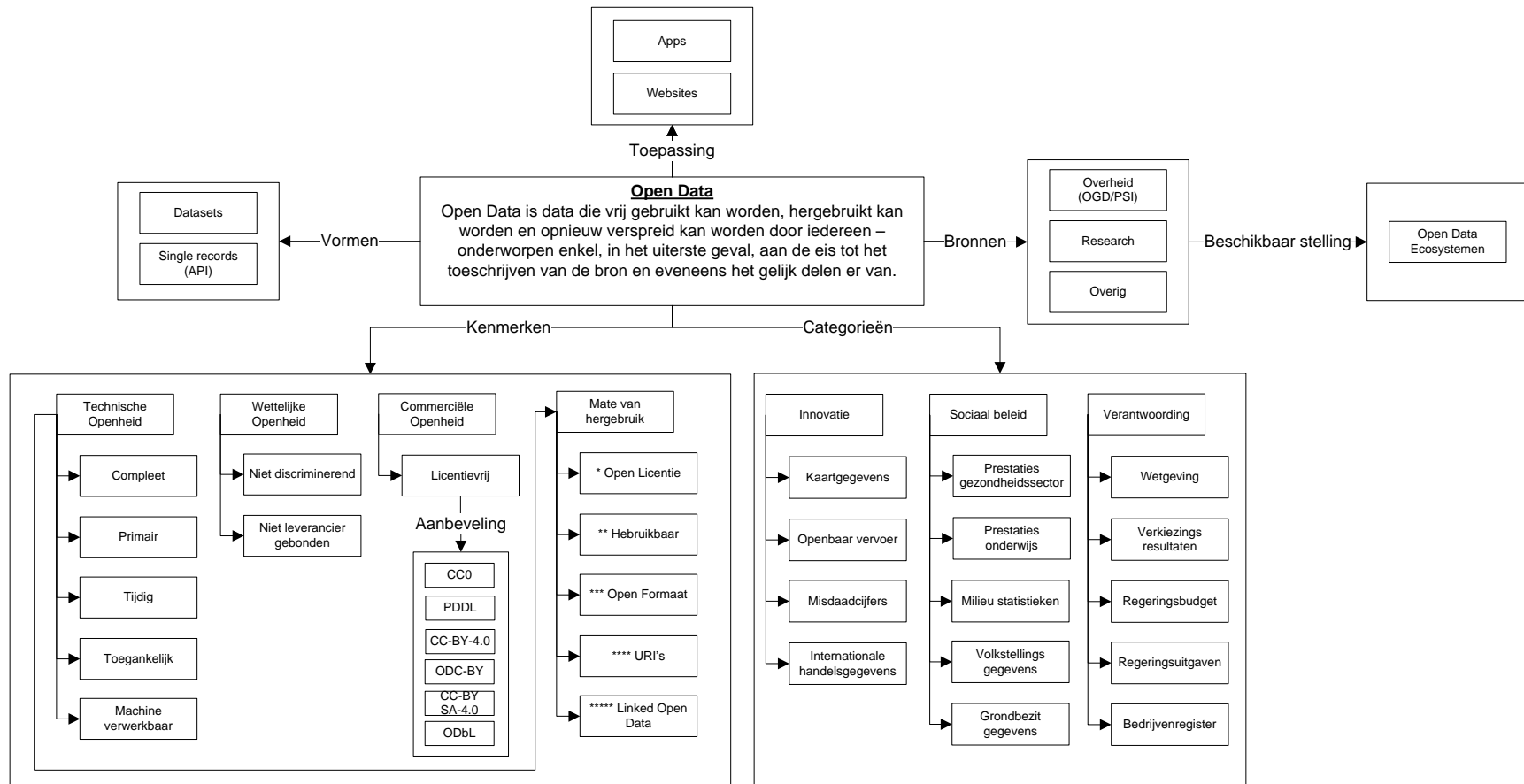
Acties naar aanleiding van antwoorden auteurs:

- Thorhildur Jetzek: Het artikel dat via de link werd meegestuurd is niet relevant voor mijn onderzoek. Het draft-artikel “Exploring the factors influencing the adoption of open government data by private organisations” geeft als conclusie dat private organisaties het belangrijk vinden dat:
 - o Om Open Data te gebruiken de inhoud en de bron van de data helder moeten zijn;
 - o De data een bruikbare licentie heeft;
 - o De continuïteit van data updates aanwezig moet zijn.Deze conclusies zijn opgesteld op basis van onderzoek bij 5 Nederlandse organisaties. Het businessmodel van deze (kleine) innovatieve organisaties is specifiek gericht op Open Data. In mijn mijlpaaldocument “Literatuurstudie” zal ik een verwijzing opnemen naar het definitieve artikel, waarbij ik zal aangeven dat er een zoektocht is gestart naar de vraagzijde van Open Data, specifiek de toepassing van Open Data in private organisaties met een Open Data business model.
- Michael Kaschesky: Geeft aan dat SpagoBI gebruik maakt van Open Data. Op de website van SpagoBI gezocht: SpagoBI is een open source aanbieder van Business Intelligence. Het gebruik van Open Data wordt in de Spago tooling beschikbaar gemaakt via CKAN (Comprehensive Knowledge Archive Network), open sourcemodel voor dataopslag en –distributie. De tooling kan (Linked) Open Data ontsluiten.
- Johann Höchtl: Het artikel is niet relevant.
- Marijn Jansen: Nogmaals gezocht met de zoekterm “Data Analytics”. Ook met die term geen relevante zoekresultaten.
- Tim Davies: Het artikel van Deloitte is niet relevant.

Bijlage F – Referentiemodel

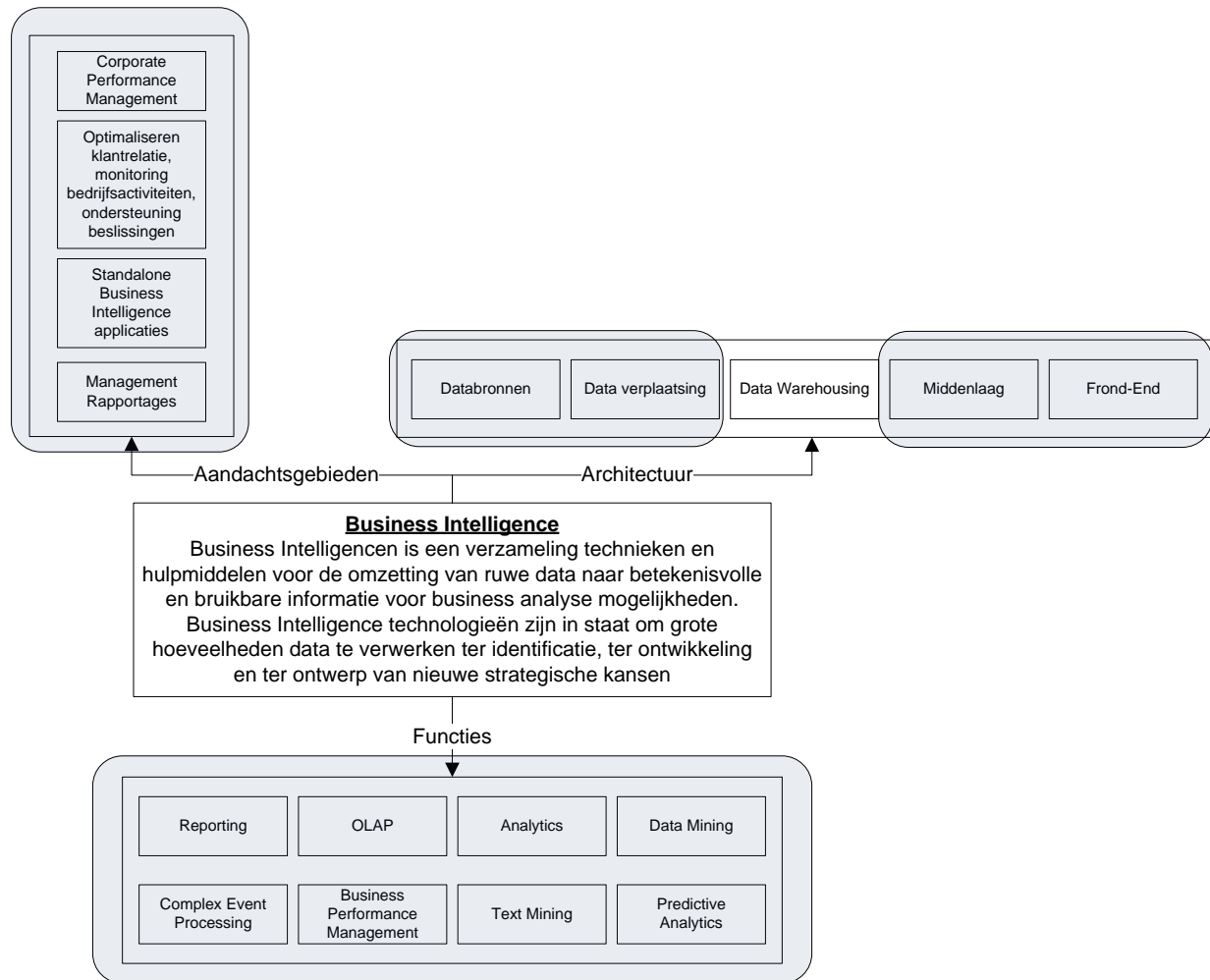
F.1. – Open Data

Onderstaande figuur toont de grafische weergave van de antwoorden die verkregen zijn op de vragen die vanuit het literatuuronderzoek zijn beantwoord. De definitie (L1), de toepassing (L4), vormen (L1), bronnen (L2), beschikbaarstelling (L2), kenmerken (L3) en categorieën (L1) van Open Data.



F.2. – Business Intelligence

Onderstaande figuur toont de grafische weergave van de antwoorden die verkregen zijn op de vragen die vanuit het literatuuronderzoek zijn beantwoord. De definitie (L5), aandachtsgebieden (L5), architectuur (L6) en functies (L7) van Business Intelligence. Het potentieel en de mogelijke problemen die Open Data met zich meenemen in de toepassing van Open Data binnen Business Intelligence is omcirkeld.



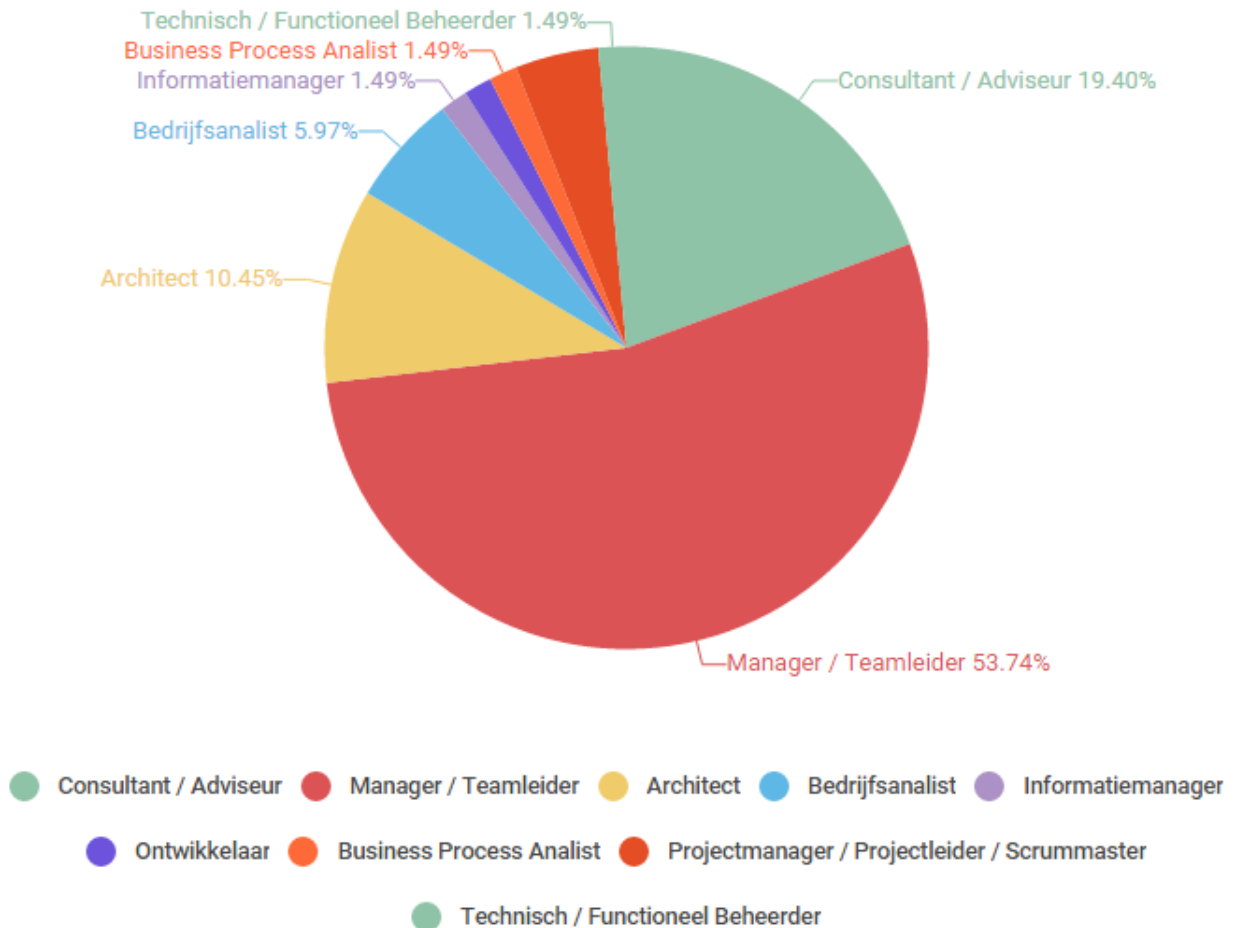
Bijlage G – Operationalisering

Op basis van het referentiemodel zijn begrippen opgenomen in een topiclijst:

- Geef de definities van Open Data en Business Intelligence
- Welke **vormen** van Open Data worden binnen de BI afdeling gebruikt?
 - o Datasets
 - o Single records (API)
- Welke **categorieën** van Open Data worden gebruikt door de BI afdeling?
 - o Innovatie → Kaartgegevens, openbaar vervoer, misdaadcijfers en internationale handelsgegevens.
 - o Sociaal beleid → Prestaties gezondheidssector, prestaties onderwijs, milieu statistieken, volkstellingsgegevens en grondbezitgegevens.
 - o Verantwoording → Wetgeving, verkiezingsresultaten, regeringsbudget, regeringsuitgaven en bedrijvenregister.
- Welke **bronnen** van Open Data worden gebruikt door de BI afdeling?
 - o Overheid
 - o Research
- Hoe wordt Open Data ontsloten door de BI-afdeling? (**Beschikbaarstelling**)
- Hoe gaat de BI-afdeling om met de **kenmerken** van Open Data?
 - o Technische openheid → Compleet, primair, tijdig, toegankelijk en machine verwerkbaar
 - o Wettelijke openheid → Niet discriminerend, niet leverancier gebonden
 - o Commerciële openheid → Licentievrij
 - o Mate van hergebruik → Vijfsterren model van Tim Berners-Lee
- Met welke **doeleinden** worden Open Data door de BI afdeling toegepast?
- Is Open Data een organisatie-**initiatief** of een initiatief vanuit de BI afdeling zelf?
- Binnen welk BI **aandachtsgebieden** worden Open Data toegepast?
 - o Corporate performance management
 - o Optimaliseren klantrelatie, monitoring bedrijfsactiviteiten, ondersteuning beslissingen
 - o Standalone BI applicaties
 - o Management rapportages
- Op welke onderdelen uit de BI **architectuur** hebben Open Data impact?
 - o Databronnen
 - o Data verplaatsing
 - o Data Warehousing
 - o Middenlaag
 - o Front-End
- Op welke BI **functies** hebben Open Data impact?
 - o Reporting
 - o OLAP
 - o Analytics
 - o Data Mining
 - o Complex Event Processing
 - o Business Performance Management
 - o Text Mining

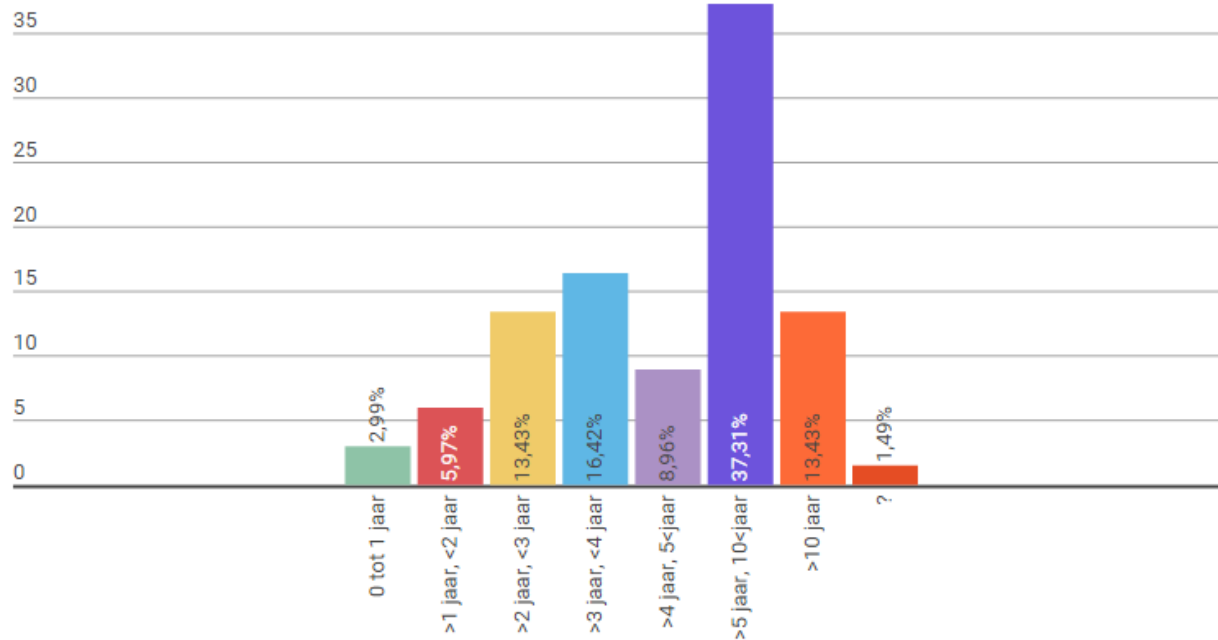
Bijlage H – Survey resultaten

H.1. ALG1 - Welke van de volgende functienamen beschrijft uw rol het beste?



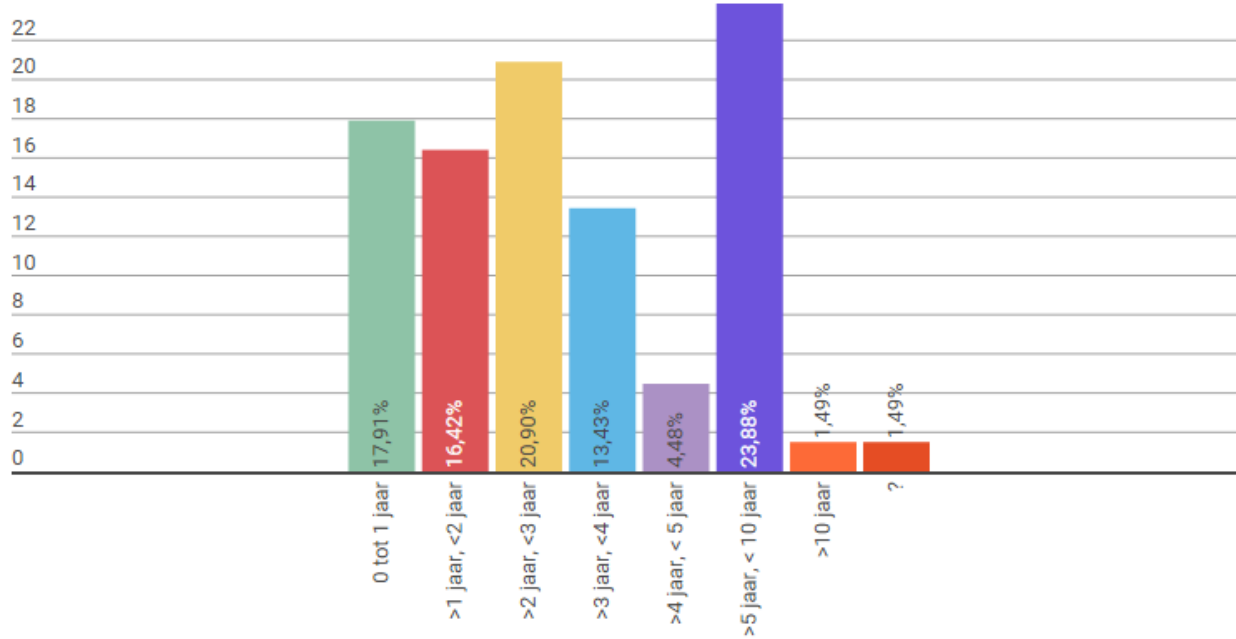
Consultant / Adviseur	13	19,40%
Manager / Teamleider	36	53,73%
Architect	7	10,45%
Bedrijfsanalist	4	5,97%
Informatiemanager	1	1,49%
Ontwikkelaar	1	1,49%
Business Process Analyst	1	1,49%
Projectmanager / Projectleider / Scrummaster	3	4,48%
Technisch / Functioneel Beheerder	1	1,49%
	67	

H.2. ALG2 - Hoe lang bestaat de Business Intelligence afdeling waar u werkzaam bent?



0 tot 1 jaar	2	2,99%
Langer dan 1 jaar, korter dan 2 jaar	4	5,97%
Langer dan 2 jaar, korter dan 3 jaar	9	13,43%
Langer dan 3 jaar, korter dan 4 jaar	11	16,42%
Langer dan 4 jaar, korter dan 5 jaar	6	8,96%
Langer dan 5 jaar, korter dan 10 jaar	25	37,31%
Langer dan 10 jaar	9	13,43%
Weet niet	1	1,49%
	67	

H.3. ALG3 - Hoe langt bent u werkzaam bij de Business Intelligence afdeling?



0 tot 1 jaar	12	17,91%
Langer dan 1 jaar, korter dan 2 jaar	11	16,42%
Langer dan 2 jaar, korter dan 3 jaar	14	20,90%
Langer dan 3 jaar, korter dan 4 jaar	9	13,43%
Langer dan 4 jaar, korter dan 5 jaar	3	4,48%
Langer dan 5 jaar, korter dan 10 jaar	16	23,88%
Langer dan 10 jaar	1	1,49%
Weet niet	1	1,49%
	67	

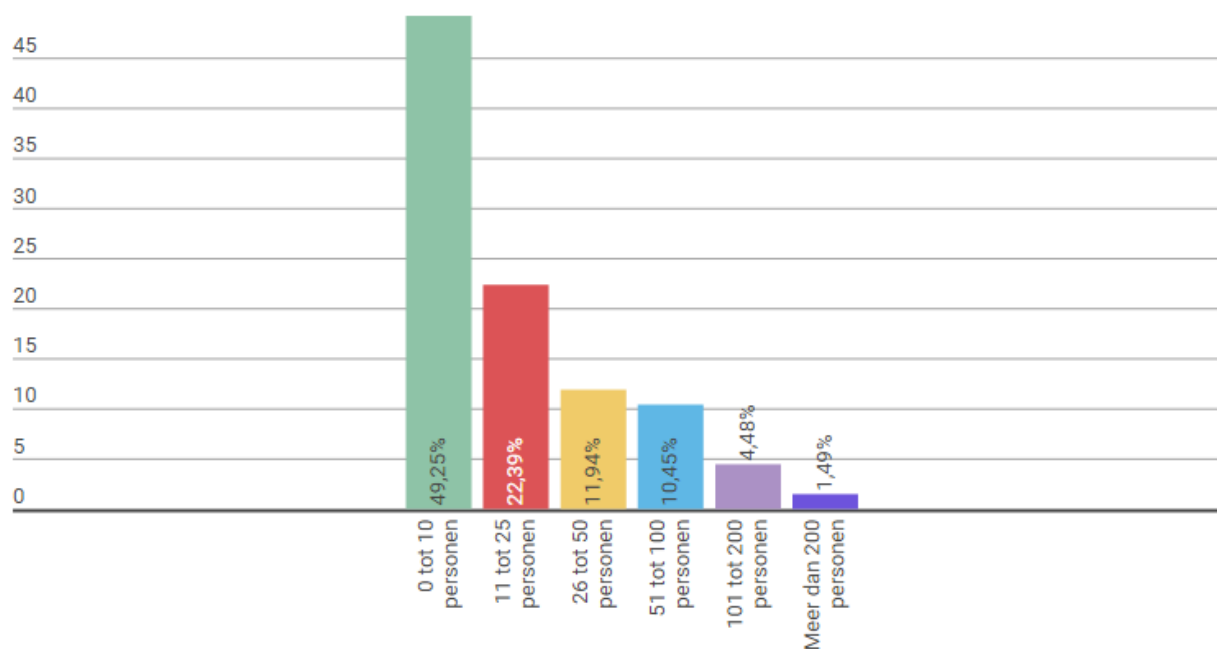
H.4. ALG4 - In welke branche is de onderneming waar u werkzaam bent actief?



Financiële instellingen	11	16,42%
Groot- en detailhandel	7	10,45%
Industrie	6	8,96%
Vervoer en opslag	6	8,96%
Adviesing, onderzoek en overige specialistische zakelijke dienstverlening	4	5,97%
Gezondheids- en welzijnszorg	4	5,97%
Informatie en communicatie	4	5,97%
Productie en distributie van en handel in elektriciteit, aardgas, stoom en gekoelde lucht	4	5,97%
Bouwnijverheid	3	4,48%
Landbouw, bosbouw en visserij	3	4,48%
Overige dienstverlening	3	4,48%
Electronica	2	2,99%
Logies-, maaltijd- en drankverstrekking	2	2,99%
Andere	1	1,49%
Automotive	1	1,49%

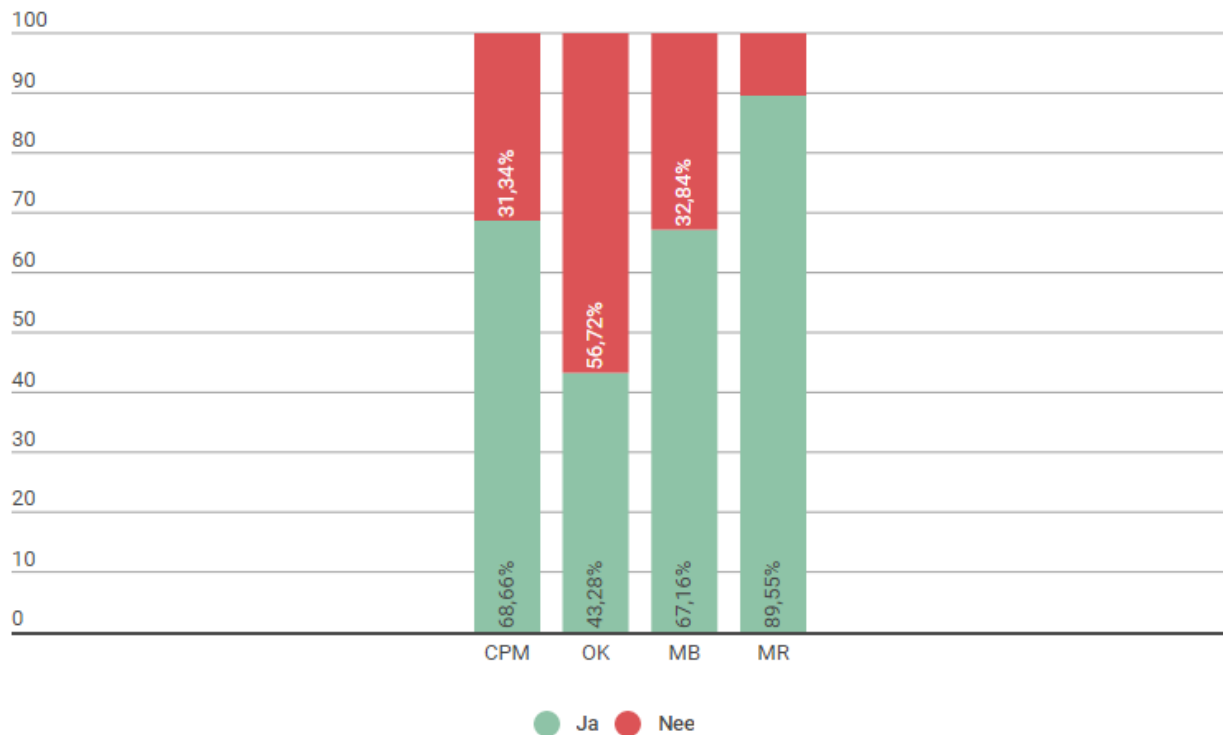
Life Science / Material Science	1	1,49%
Loterijen	1	1,49%
Toerisme	1	1,49%
Vastgoed	1	1,49%
Verhuur van en handel in onroerend goed	1	1,49%
Vleesproductie	1	1,49%
	67	

H.5. BIA1 - Hoeveel personen zijn werkzaam bij de Business Intelligence afdeling?



0 tot 10 personen	33	49,25%
11 tot 25 personen	15	22,39%
26 tot 50 personen	8	11,94%
51 tot 100 personen	7	10,45%
101 tot 200 personen	3	4,48%
Meer dan 200 personen	1	1,49%
	67	

H.6. BIA2 - Voor welke doeleinden wordt Business Intelligence binnen de onderneming ingezet?



	CPM Toepassing	OK Toepassing	MB Toepassing	MR Toepassing
Ja	68,66%	43,28%	67,16%	89,55%
Nee	31,34%	56,72%	32,84%	10,45%

Percentages gaan over 67 respondenten

CPM = Corporate Performance Management

OK = Optimaliseren Klantrelatie

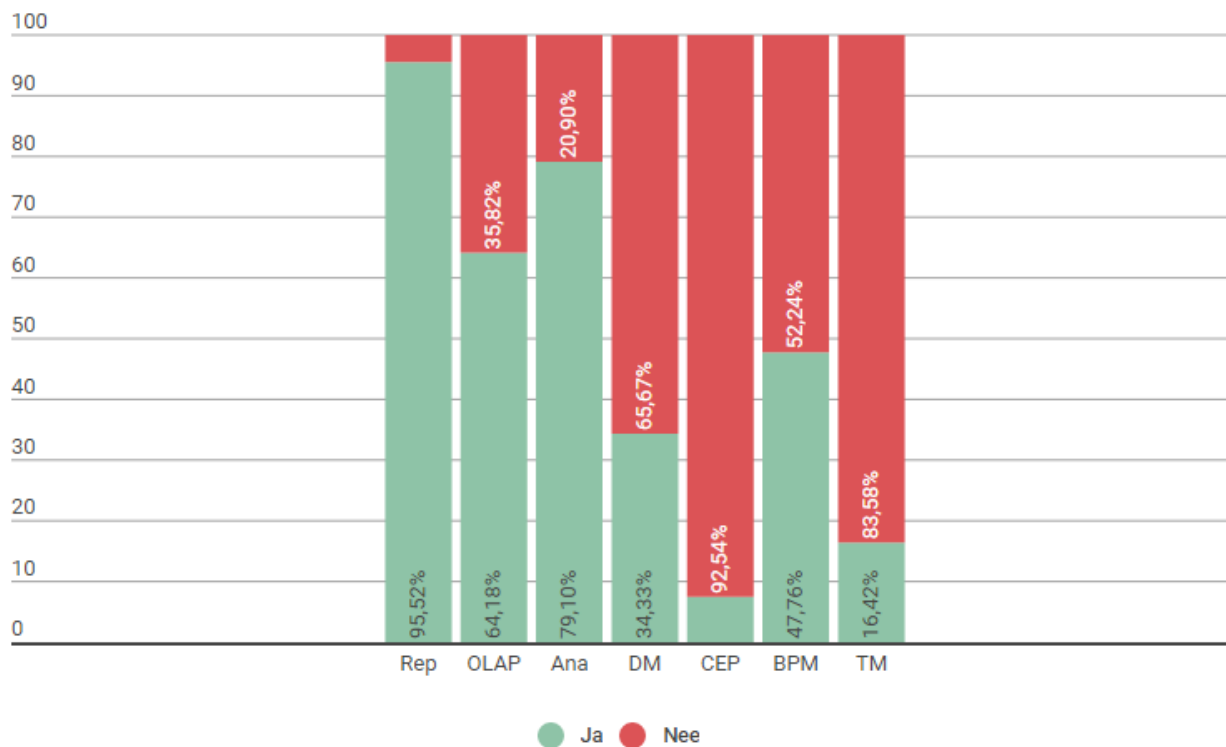
MB = Monitoring Bedrijfsactiviteiten

MR = Management Rapportages

Overig genoemde doeleinden BI:

- Data hub naar andere systemen
- Doelgroepanalyses, fraudedetectie en retentiebeleid
- Externe dienstverlening
- Future prediction
- Innovatie en procesoptimalisatie
- Marketing
- Optimaliseren zorginkoop
- Proof of concept

H.7. BIA3 - Welke Business Intelligence functies worden ingezet?



	Rep	OLAP	Ana	DM	CEP	BPM	TM
Ja	95,52%	64,18%	79,10%	34,33%	7,46%	47,76%	16,42%
Nee	4,48%	35,82%	20,90%	65,67%	92,54%	52,24%	83,58%

Percentages gaan over 67 respondenten

Rep = Reporting

OLAP = OnLine Analytical Processing

Ana = Analytics

DM = Data Mining

CEP = Complex Event Processing

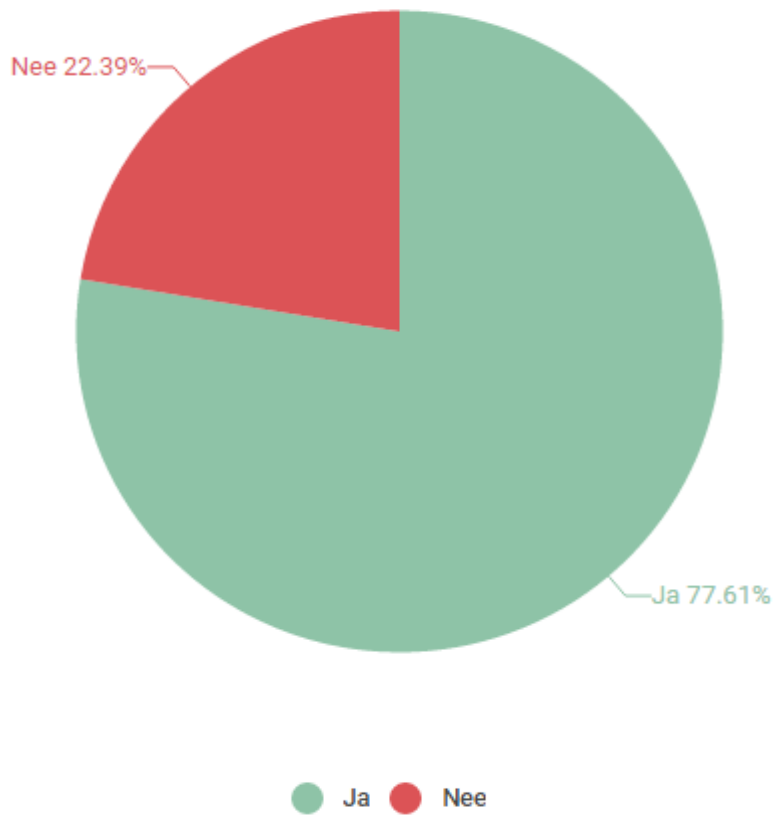
BPM = Business Performance Management

TM = Text Mining

Overig genoemde functies BI:

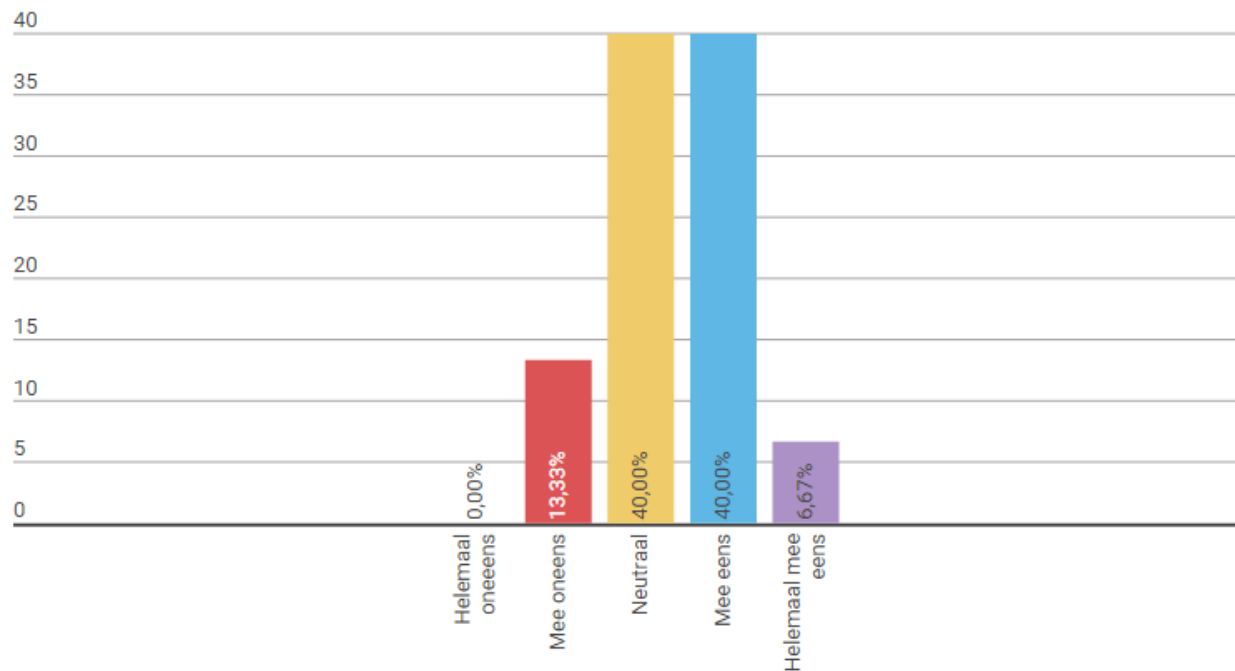
- GEO
- Big data
- Onderzoek

H.8. ODB1 - Is de Business Intelligence afdeling bekend met het fenomeen Open Data?



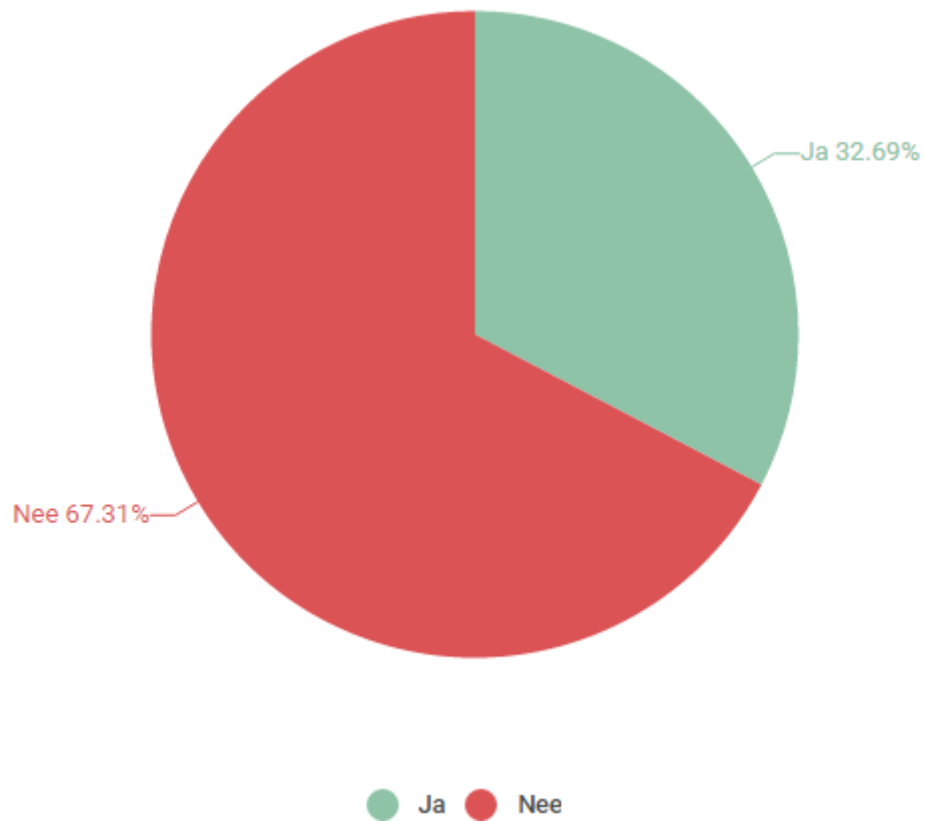
Ja	52	77,61%
Nee	15	22,39%
	67	

H.9. ODB2 - De toepassing van Open Data zou voor de afdeling Business Intelligence waar ik werkzaam ben van toegevoegde waarde kunnen zijn?



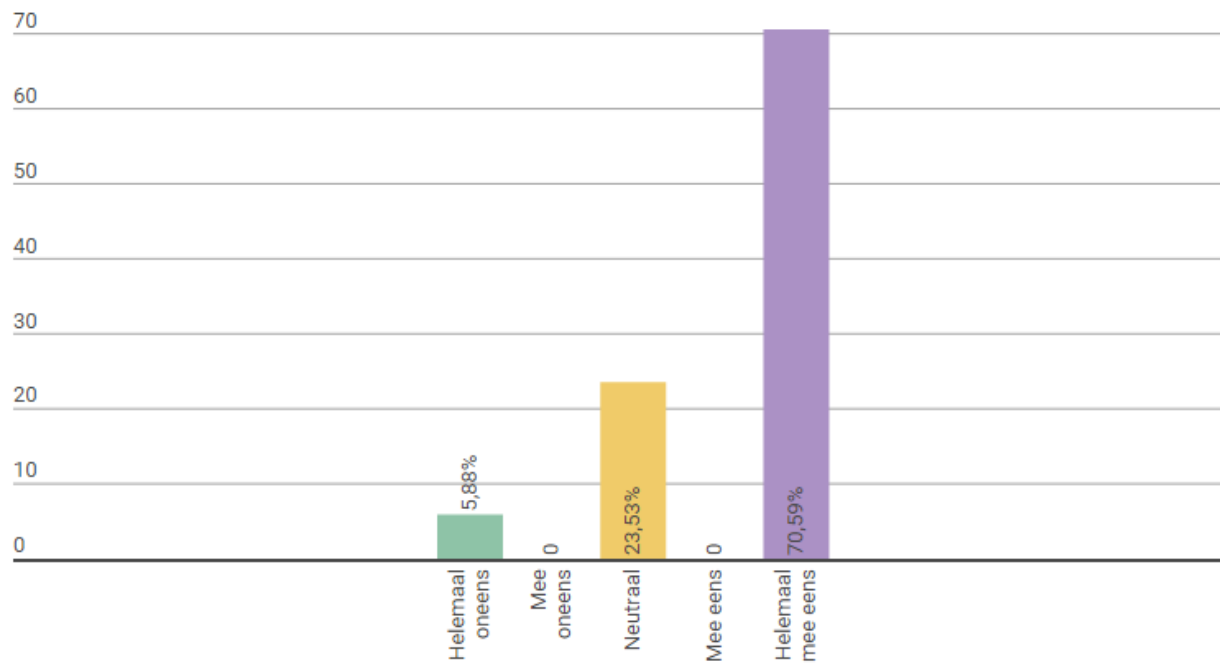
Helemaal mee oneens	0	0,00%
Mee oneens	2	13,33%
Neutraal	6	40,00%
Mee eens	6	40,00%
Helemaal mee eens	1	6,67%
	15	

H.10. ODG1 - Maakt de Business Intelligence afdeling gebruik van Open Data?



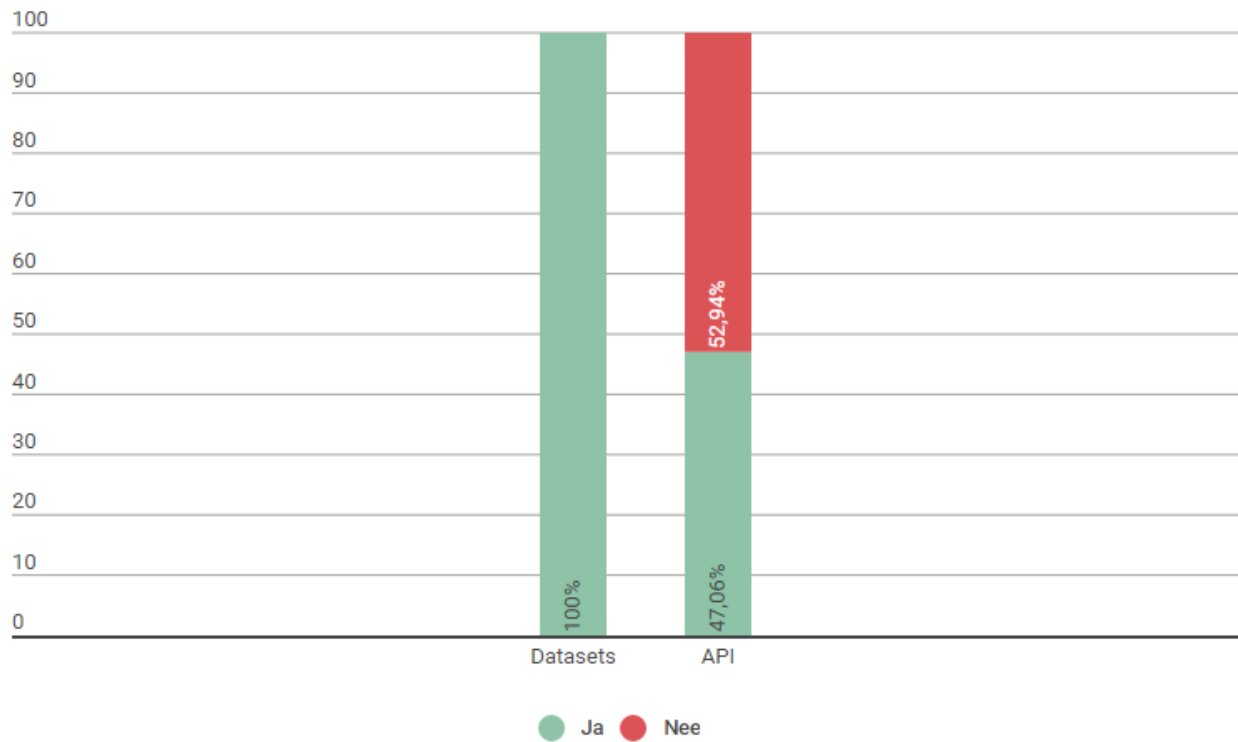
Ja	17	32,69%
Nee	35	67,31%
	52	

H.11. ODB3 - De toepassing van Open Data is van toegevoegde waarde voor de afdeling Business Intelligence?



Helemaal mee oneens	1	5,88%
Eens	4	23,53%
Helemaal mee eens	12	70,59%
	17	

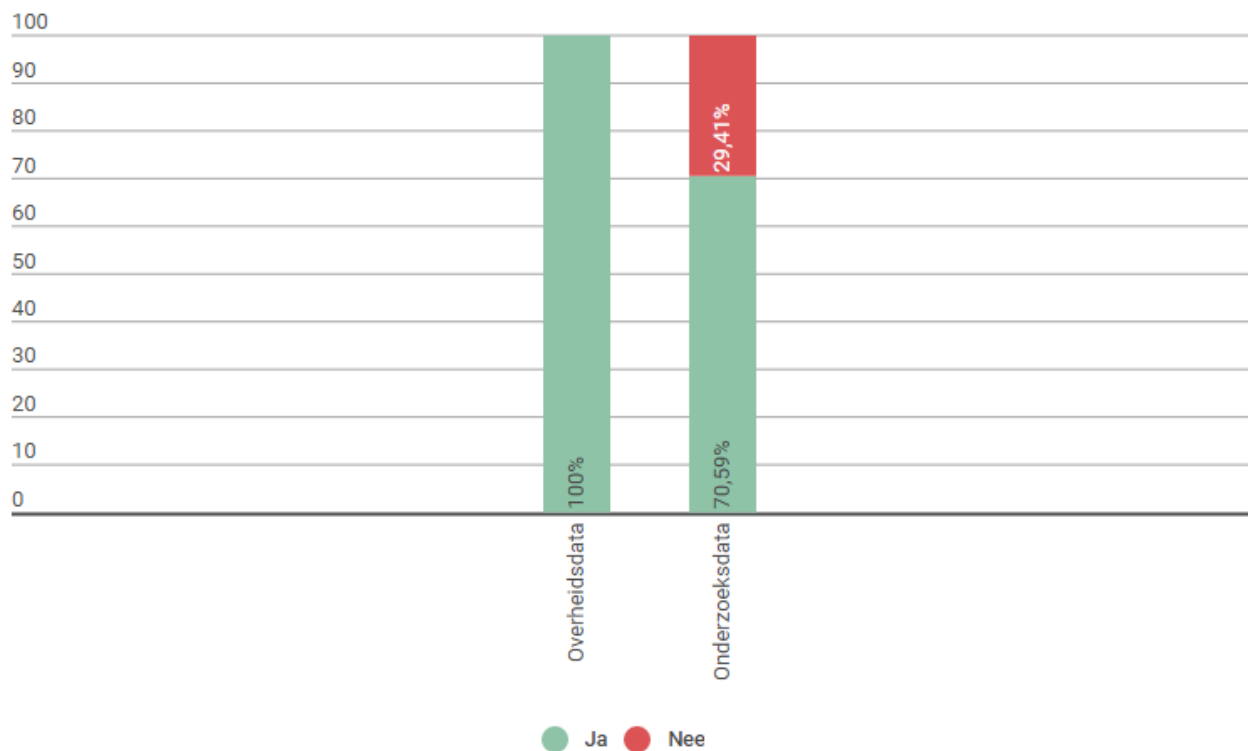
H.12. ODG2 - Van welke vormen van Open Data wordt op de Business Intelligence afdeling waar u werkzaam bent gebruik gemaakt?



	Datasets	API
Ja	100%	47,06%
Nee	0%	52,94%

Percentages gaan over 17 respondenten

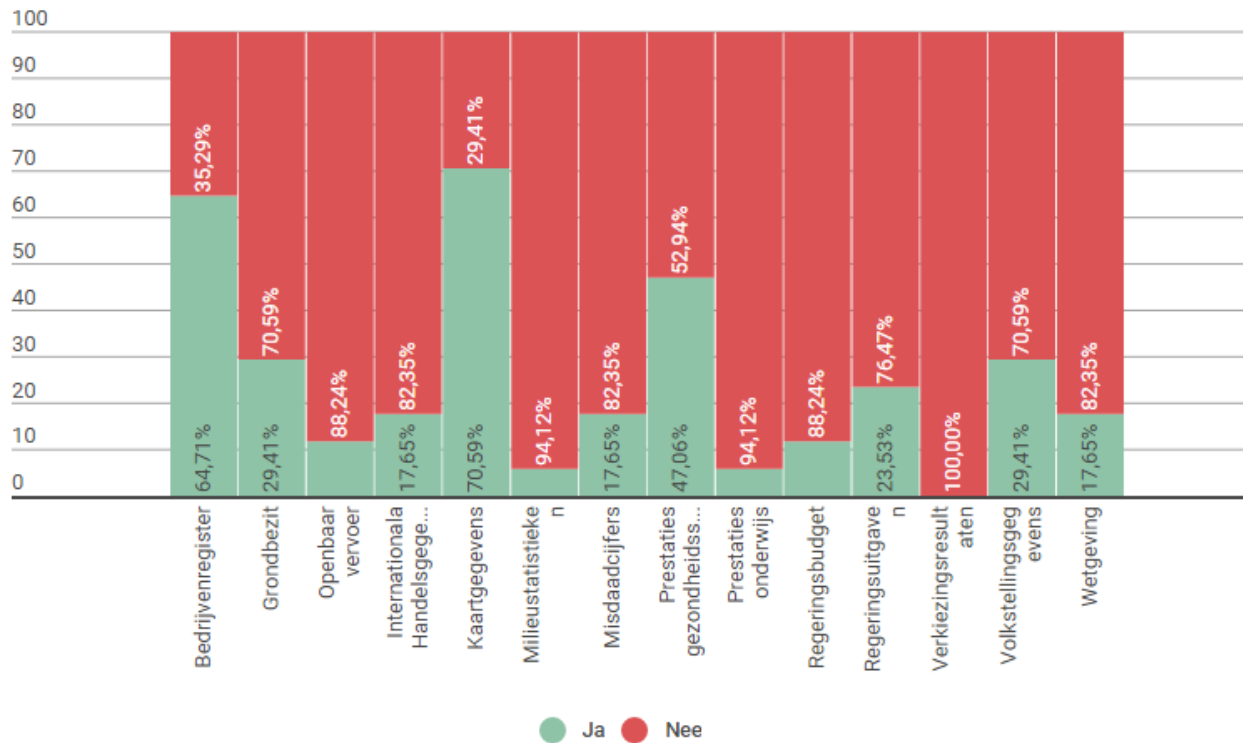
H.13. ODG3 – Van welke bronnen van Open Data wordt door de Business Intelligence afdeing gebruik gemaakt?



	Overheidsdata	Onderzoeksdata
Ja	100%	70,59%
Nee	0%	29,41%

Percentages gaan over 17 respondenten

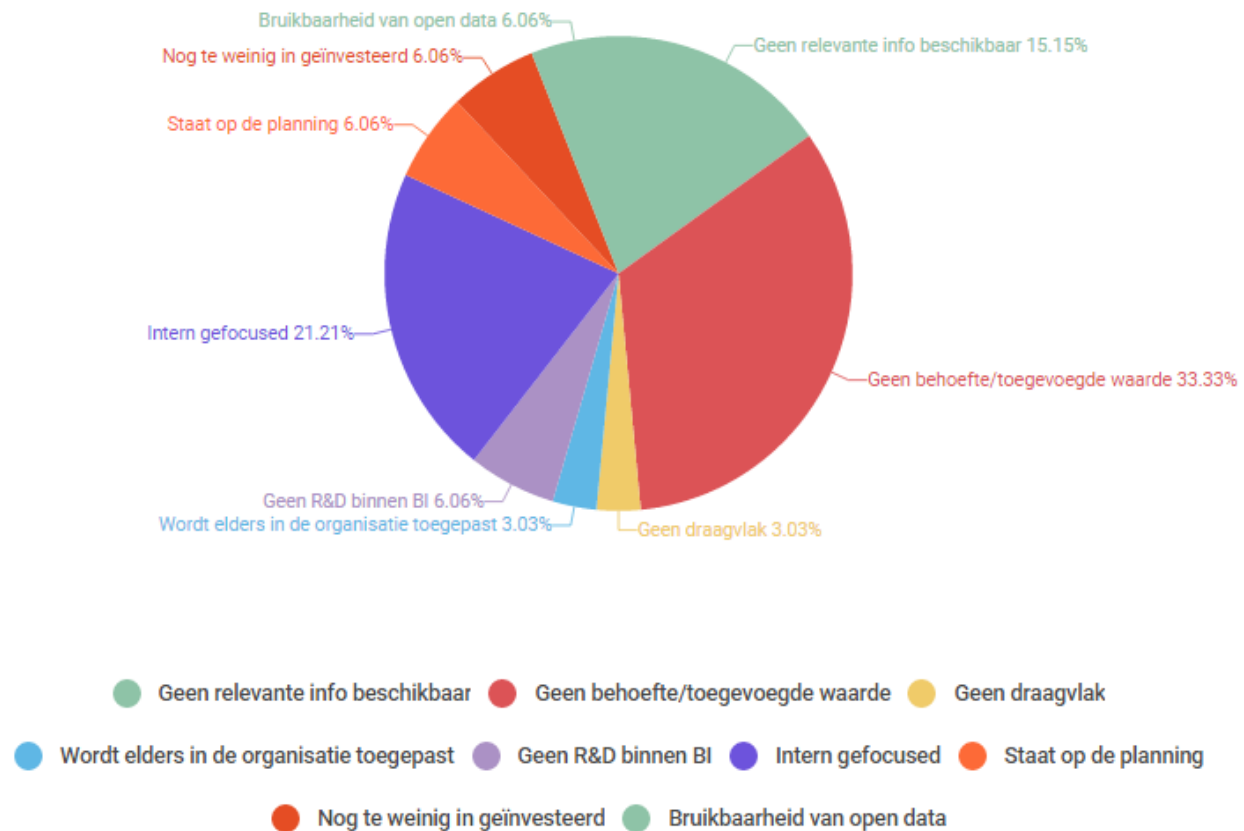
H.14. ODG4 - Van welke categorieën van Open Data wordt door uw Business Intelligence afdeling gebruik gemaakt?



	Ja	Nee
Bedrijvenregister	64,71%	35,29%
Grondbezit	29,41%	70,59%
Openbaar vervoer	11,76%	88,24%
Internationale Handelsgegevens	17,65%	82,35%
Kaartgegevens	70,59%	29,41%
Milieustatistieken	5,88%	94,12%
Misdaadcijfers	17,65%	82,35%
Prestaties gezondheidssector	47,06%	52,94%
Prestaties onderwijs	5,88%	94,12%
Regeringsbudget	11,76%	88,24%
Regeringsuitgaven	23,53%	76,47%
Verkiezingsresultaten	0,00%	100,00%
Volkstellingsgegevens	29,41%	70,59%
Wetgeving	17,65%	82,35%

Percentages gaan over 17 respondenten

H.15. ODG5 - Waarom maakt de Business Intelligence afdeling geen gebruik van Open Data?



Geen relevante info beschikbaar	5	15,15%
Geen behoefte/toegevoegde waarde	11	33,33%
Geen draagvlak	1	3,03%
Wordt elders in de organisatie toegepast	1	3,03%
Geen R&D binnen BI	2	6,06%
Intern gefocused	7	21,21%
Staat op de planning	2	6,06%
Nog te weinig in geïnvesteerd	2	6,06%
Bruikbaarheid van open data	2	6,06%
	33	

Bijlage I – Overzicht documenten documentatiestudie

De volgende documenten zijn bij de documentatiestudie geanalyseerd. Het was niet mogelijk om de totstandkoming en de doelgroep van de documenten boven water te krijgen.

ORGANISATIE	OMSCHRIJVING	DOEL
Achmea	BvR_AchmeaVanBovenMedewerkers_A3	Voorbeeld toepassing Open Data
Achmea	BvR_FO VdK 20140813	Functioneel Ontwerp. Open Data toepassing wordt beschreven
Achmea	BvR_Gebruikte Externe Bronnen	Gebruikte Open Data bronnen
Achmea	BvR_Geo_SenI_ACO_update	Overview GEO landschap
Achmea	BvR_PaviljoenGegevens	Toelichting stelsel basisregistratie
Achmea	BvR_RWSongevallenHecto20150225_A58landscape	Voorbeeld toepassing Open Data
Achmea	BvR_VeiligeRoutes20150507	Voorbeeld toepassing Open Data
Achmea	GK_Achmea Data Glossary_ Externe Bronnen	Gebruikte Open Data bronnen
Achmea	GK_PSA Klantsignalen rapportage	Project start architectuur, inclusief Open Data toepassing
Achmea	KS_PredictingBurglary_DataFields20150310GlobalEstimate	Voorbeeld toepassing Open Data
Achmea	LV_NieuwRisicomodelAuto v3	Voorbeeld toepassing Open Data
Enexis	GP_organogram Enexis	Voorbeeld toepassing Open Data
Enexis	JH_GAS Priotool Datastromen	Voorbeeld toepassing Open Data
Enexis	JH_Thematische kaart	Voorbeeld toepassing Open Data
Enexis	JH_Thematische kaart2	Voorbeeld toepassing Open Data
Enexis	RS_ACO - Presentatie Open Data 02-10-2014 in De Croon	Open data en invloed Enexis
Enexis	RS_BI architectuur	Indeling BI architectuur binnen Enexis
Enexis	RS_DataHub	Opbouw datahub
Enexis	RS_Open Data agenda 2015-2018v0 99	Open data agenda
Enexis	RS_TEI Userstories Gasleidingen	Voorbeeld case Open Data
Enexis	RS_Type_BI_Selfservice	Opbouw BI Selfservice

Bijlage J – Boomstructuur NVivo

Name	Sources	References
Organisatie	0	0
Achmea	0	0
Interview XXX - 12082015	1	1
Interview XXX - 12082015	1	1
Interview XXX - 13072015	1	1
Interview XXX - 14072015	1	1
Interview XXX - 07092015	1	1
Interview XXX - 07092015	1	1
Interview XXX - 14072015	1	1
Afdeling	1	2
Enexis	0	0
Interview XXX - 15092015	1	1
Interview XXX - 09092015	1	1
Interview XXX - 09092015	1	1
Interview XXX - 09092015	1	1
Interview XXX - 09092015	1	1
Interview XXX - 15092015	1	1
Anonieme Organisatie	0	0
Interview XXX - 24082015	1	1
Open Data	0	0
Vormen	0	0
Single records	5	5
Datasets	8	8
Categorieën	0	0
Innovatie	0	0
Openbaar vervoer	1	1
Misdaadcijfers	2	2
Internationale handelsgegevens	2	2
Kaartgegevens	9	10
Sociaal Beleid	0	0

Prestaties gezondheidssector	0	0
Prestaties onderwijs	2	2
Milieu statistieken	3	3
Volkstellingsgegevens	5	5
Grondbezit gegevens	5	5
Verantwoording	0	0
Verkiezingsresultaten	0	0
Regeringsbudget	0	0
Regeringsuitgaven	0	0
Bedrijvenregister	3	3
Wetgeving	5	5
Kenmerken	0	0
Technisch	0	0
Toegankelijk	6	6
Compleet	8	8
Machine verwerkbaar	8	8
Tijdig	9	9
Primair	8	10
Wettelijk	0	0
Licentievrij	9	9
Commercieel	0	0
Niet gebonden	8	8
Niet discriminerend	9	9
Linked Open Data	1	3
Doel gebruik en-of ontsluiting	0	0
Toegevoegde waarde	5	5
Doel	10	16
Toepassing	0	0
Ad hoc	4	6
Structureel	3	6
Bronnen	1	1
Onderzoek	2	2
Overheid	6	7
Positieve kant	1	3

Ontsluiting	1	1
Frequentie	0	0
Periodiek	1	1
Eenmalig	2	2
Wijze	1	1
Automatisch	6	6
Handmatig	8	9
Publicatie	1	1
Tooling	1	1
Strategie	1	1
Volwassenheidsniveau in gebruik	4	4
Impact	5	7
BI Functies	5	5
BI aandachtsgebieden	3	6
Negatief	1	2
BI Architectuur	6	7
Forecasting	1	1
Gebruik	8	9
Bij Big Data	1	1
Definitie	11	11
Initiatief	8	11
Veelgenoemde open data sets	10	20
Problemen	8	20
BI	1	1
Positionering organisatie	0	0
Subteam	12	16
Aandachtsgebieden	0	0
GEO	1	1
Standalone BI	2	2
Business Process Monitoring	2	2
Prestatiemanagement (CPM)	4	4
Optimaliseren Klantrelatie	4	5
Management Rapportages	5	5
Functies	2	2

Process Mining		1	1
Geografische analyse		1	1
Complex Event Processing		2	2
Reporting		3	3
OLAP		3	3
Analytics		3	3
Data Mining		3	3
Business Performance Management		3	3
Text Mining		3	3
Aantal medewerkers		5	5
Architectuur		5	6